

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 8, 2003, 14:36:31 ; Search time 2678 Seconds  
(without alignments)  
12143.162 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcgcagtgtccga.....ggggcacgacataggggt 1338

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rtd:\*
- 26: em\_gss\_pbg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

pred, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	980.2	73.3	1201	13	BX352919
2	889.2	66.5	1201	9	AL519338
3	884	66.1	1086	12	BM548059
4	880.6	65.8	977	9	AL526237

5	832.2	62.2	1080	12	BM923064
6	794.6	59.4	873	9	AL525111
7	786.4	58.8	873	12	BI837148
8	763.8	57.1	773	12	BI827333
9	753	56.3	856	12	BM014974
10	737.6	55.1	794	9	AU141625
11	737.2	55.1	964	13	BX370424
12	704.8	52.7	884	13	BU181305
13	701.2	52.4	1145	12	BM473687
14	697.4	52.1	932	13	BQ436653
15	688.2	51.4	734	12	BI253559
16	683.2	51.1	770	9	AL522085
17	660	49.3	741	12	BM010281
18	627.6	46.9	773	12	BI818555
19	622.4	46.5	664	9	AI963473
20	620.4	46.4	679	10	BG393676
21	617	46.1	895	10	BF797243
22	614.4	45.9	1036	12	BM45513
23	601.4	44.9	890	13	BU172522
24	589.8	44.1	1024	12	BM465183
25	580.4	43.4	590	10	BG282890
26	569.2	42.5	1271	11	AK002912
27	568.8	42.4	1201	9	AL556719
28	567.8	42.4	823	12	BM019985
29	558.4	41.7	560	9	AU280528
30	557.4	41.7	595	10	BE621964
31	553.6	41.4	902	14	CB182842
32	552.8	41.3	1132	12	BM424026
33	539.6	40.3	996	14	BY703409
34	535.8	40.0	1265	11	AK013214
35	535	40.0	803	13	BQ571032
36	532	39.8	744	10	BE267359
37	532	39.8	902	13	BQ964436
38	527.6	39.4	788	12	BI254417
39	525.6	39.3	874	12	BI415573
40	520.8	38.9	1348	11	AK004291
41	520.6	38.9	737	12	BM019007
42	517.4	38.7	670	10	BE563597
43	510.6	38.2	869	12	BI329060
44	505.2	37.8	602	9	AL706316
45	504.4	37.7	1178	12	BM461077

ALIGNMENTS

RESULT 1  
BX352919  
LOCUS  
DEFINITION BX352919 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens  
CDNA clone CS0DB006YC20 5-PRIME, mRNA sequence.  
ACCESSION BX352919  
VERSION BX352919.1 GI:30347710  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Li,W.B., Gruber,C., Jessee,J., and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DB006B100P1&cluster=4390.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

BX352919 1201 bp mRNA linear EST 02-MAY-2003  
Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens  
CDNA clone CS0DB006YC20 5-PRIME, mRNA sequence.  
BX352919.1 GI:30347710  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li,W.B., Gruber,C., Jessee,J., and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DB006B100P1&cluster=4390.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

BM923064	AGENCOURT
AL525111	AL525111
BI837148	603089920
BI827333	603077811
BM014974	603640923
AU141625	AU141625
BX370424	BX370424
BU181305	AGENCOURT
BM473687	AGENCOURT
BQ436653	AGENCOURT
BI253559	602973485
AL522085	AL522085
BM010281	603631154
BI818555	603033078
AI963473	wt63q02.x
BG393676	602412094
BF797243	60257745
BM45513	AGENCOURT
BU172522	AGENCOURT
BM465183	AGENCOURT
BG282890	602406016
AK002912	Mus muscu
AL556719	AL556719
BM019985	603648507
AU280528	AU280528
BE621964	601440609
CB182842	AGENCOURT
BM424026	AGENCOURT
BY703409	BY703409
AK013214	Mus muscu
BQ571032	UI-M-FB0-
BE267359	601189568
BQ964436	AGENCOURT
BI254417	602976749
BI415573	602988350
AK004291	Mus muscu
BM019007	603646981
BE563597	601334712
BI329060	602980475
AL706316	DKFP686C
BM461077	AGENCOURT

TAAAGGCCCCCGTGGAGAGGCTGTGCCCCACCTTGGCCCGGCCCCAAGAGGAGCTCTAGGCGAGAAAGGTATAGGCG

[illegible]

QY  
03 CATGGCGGCGGCGGCCCATTCGCCTCAGGGGCCCCATCACAGAACACAGGAAAGAGTGGGAAGGCTCAGCCACAT 142



Db 445 GATGGCGGGGGCCGACAGTGGGAGACGAGAGACCAACCTCTCTGACGATCTGCTGGA 504  
QY 500 GGCCTGTGGGCGGTGTGAGGAGTGCCTGAAGCCTACGTGCGACATCCACACTGGCTCAG 559  
Db 505 GGCCTGTGGGCGGTGTGAGGAGTGCCTGAAGCCTACGTGCGACATCCACACTGGCTCAG 564  
QY 560 TGGCAGTGGGCGGTGTGAGTGTGATCTCTCGAGGCGCTGGCTGAAGAGGCGGTCAA 619  
Db 565 TGGCAGTGGGCGGTGTGAGTGTGATCTCTCGAGGCGCTGGCTGAAGAGGCGGTCAA 624  
QY 620 GATGGCATGCCAGAGCCTGGCCACACCGCATCGTGGCCAGACACCTGCTGGGACGGC 679  
Db 625 GATGGCATGCCAGAGCCTGGCCACACCGCATCGTGGCCAGACACCTGCTGGGACGGC 684  
QY 680 CAAGATGCTGTGACAGAGGCGCAACACCCAGCCAGCTGCGTCTAGAGTGTGACACC 739  
Db 685 CAAGATGCTGTGACAGAGGCGCAACACCCAGCCAGCTGCGTCTAGAGTGTGACACC 744  
QY 740 GGGTGGCACCAACCATCTATGACTCCACGCGCTGGAGCAGGCGGCTGCGAGCAGCCAC 799  
Db 745 GGGTGGCACCAACCATCTATGACTCCACGCGCTGGAGCAGGCGGCTGCGAGCAGCCAC 804  
QY 800 CATGAGCGCGCTGAGAGTGCACACTCCCGGCGCAAGAGACTCAGCAGAAAGTA-GGCTG 858  
Db 805 CATGAGCGCGCTGAGAGTGCACACTCCCGGCGCAAGAGACTCAGCAGAAAGTAGGGCTG 864  
QY 859 GGCTGTGGCCAT-CCTTTCTGCGCTGTGCGCCCTGCTCTCCCTGT-GTCCCTTCCCT 916  
Db 865 GGCTGTGGCCAT-CCTTTCTGCGCTGTGCGCCCTGCTCTCCCTGT-GTCCCTTCCCT 924  
QY 917 GAGGACTGC-GGCTCCCTCCCTCTGCTGATGAGG-TCCTCTACTGCTCTCTCCCTT 973  
Db 925 GAGGACTCGGGTCCCTCTCTCTGATGAGGCGGCTCTCTCTCTCTCTCTCTCTCT 984  
QY 974 GCACAGGAAA 984  
Db 985 GCACAGGAAA 995

RESULT 4  
AL526237  
LOCUS AL526237 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED EST 23-MAY-2003  
DEFINITION cDNA clone CS0DC016VG13 5-PRIME, mRNA sequence.  
ACCESSION AL526237  
VERSION AL526237.2 GI:31064098  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 977)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT On Feb 13, 2001 this sequence version replaced gi:12789730.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
Bp 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC016AD070P1&cluster=4390.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DC016AD070P1.  
Location/Qualifiers  
1..977

FEATURES  
Source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clone="CS0DC016VG13"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 175 a 312 c 316 g 173 t 1 others  
ORIGIN

Query Match 65.8%; Score 880.6; DB 9; Length 977;  
Best Local Similarity 99.0%; Pred. No. 3e-168;  
Matches 916; Conservative 1; Mismatches 5; Indels 3; Gaps 3;  
QY 7 CGCAGTCTGTCCAGGACCAACAGATGGCAGCTGCGGAGCGGTCTCCGCGCGCGTGGCT 66  
Db 56 CCCGGGATGTCCAGGAGCAACAGATGGCAGCTGCGGA-CCGTCTCCGCGCGCGTGGCT 114  
QY 67 TCGTGGCGGGGCGCATGCGGGGCGCATGCGGAGGGCCCTCATCAGAGCAGGAAAAG 126  
Db 115 TCGTGGCGGGGCGCATGCGGGGCGCATGCGGAGGGCCCTCATCAGAGCAGGAAAAG 174  
QY 127 TGGAGCTCAGCAGTACTGCGCAGTGCACCAACACAGAGAACCTATCTCACTTCAAG 186  
Db 175 TGAAGCTCAGCAGTACTGCGCAGTGCACCAACACAGAGAACCTATCTCACTTCAAG 234  
QY 187 CTCTGCGCTTGGCGGACCAACAGTCCAAACAGAGGTGCTGAGAGCTGCTCTGCTGCTCA 246  
Db 235 CTCTGCGCTTGGCGGACCAACAGTCCAAACAGAGGTGCTGAGAGCTGCTCTGCTGCTCA 294  
QY 247 TCTTTGGCACAGAGCTCATGCTGCGAGCTCTCTGCGAGAGGTGGCTCTCTGCTGCTCA 306  
Db 295 TCTTTGGCACAGAGCTCATGCTGCGAGCTCTCTGCGAGAGGTGGCTCTCTGCTGCTCA 354  
QY 307 CCACCTGAACACATCTTGGTCTCGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366  
Db 355 CCACCTGAACACATCTTGGTCTCGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414  
QY 367 TGCTGCCCCCAACACACAGGGTGTCTGGGGTCTTGGCCCAACCTGCTGCTGCTGCTGCTG 426  
Db 415 TGCTGCCCCCAACACACAGGGTGTCTGGGGTCTTGGCCCAACCTGCTGCTGCTGCTGCTG 474  
QY 427 AAGGGGCATAGTATGATGGCGGGGGCGCCACACGTGGGAGCAGCAGAGCAAGCTCTG 486  
Db 475 AAGGGGCATAGTATGATGGCGGGGGCGCCACACGTGGGAGCAGCAGAGCAAGCTCTG 534  
QY 487 AGCATCTGTGAGGCGCTGTGGGCGGTGTGAGAGGTGCTGAAGCTTACGTGCGATCC 546  
Db 535 AGCATCTGTGAGGCGCTGTGGGCGGTGTGAGAGGTGCTGAAGCTTACGTGCGATCC 594  
QY 547 ACACCTGGCCTCAGTGGCAGTGGGCTTCCGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 606  
Db 595 ACACCTGGCCTCAGTGGCAGTGGGCTTCCGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 654  
QY 607 AAGGAGCGGTCAAGATGGGATGCCAGAGCTGGCCACCGCATCGCTGCTGCTGCTGCTGCTG 666  
Db 655 AAGGAGCGGTCAAGATGGGATGCCAGAGCTGGCCACCGCATCGCTGCTGCTGCTGCTGCTG 714  
QY 667 TGCTGGGAGCGGCCAAGATGCTGTCACAGGGGCAACACCGAGCTGCTGCTGCTGCTGCTG 726  
Db 715 TGCTGGGAGCGGCCAAGATGCTGTCACAGGGGCAACACCGAGCTGCTGCTGCTGCTGCTG 774  
QY 727 ACCTGTGACCCCGGGTGGCACCACCATCATCATGAGTCCACGCTTGGAGAGGGGGGGC 786  
Db 775 ACCTGTGACCCCGGGTGGCACCACCATCATCATGAGTCCACGCTTGGAGAGGGGGGGC 834  
QY 787 TGGCAGCAGCACCATCAGCGCGCTGCGAGCTGCCACCTGCGGGGCAAGAGAGCTCAGCA 846  
Db 835 TGGCAGCAGCACCATCAGCGCGCTGCGAGCTGCCACCTGCGGGGCAAGAGAGCTCAGCA 893  
QY 847 GAAAGTAGGCTGGGCTCTGGCCATTCCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906  
Db 894 GAAAGTAGGCTGGGCTCTGGCCATTCCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 953



**Qy** 907 CCCTTCCCCCTGAGGACTGGGGCTCC 931  
||||| ||||| ||||| ||||| |||||  
**Db** 954 CCCTT-CCCTGAAGACTGGGGCTCC 977

RESULT 5	1080 bp	linear	EST 12-MAR-2002
BM923064		mRNA	
LOCUS	1080 bp		
DEFINITION	AGSCNCOURT_6631985 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5757031 5', mRNA sequence.		

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1080)  
NTH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.

BASE COUNT 189 a 358 c 350 g 183 t

ORIGIN

Query Match 62.2%; Score 832.2; DB 12; Length 1080;

Best Local Similarity 91.7%; Pred. No. 2e-158;

Matches 973; Conservative 0; Mismatches 68; Indels 20; Gaps 8

Qy	75	GGGGCCGCATGGCGGGGGCCATCGCGAGGGCCTCATCAGACAGGAAAAGTGGAAAGCT	134
Db	74	GGGGCCGCATGGCGGGGGCCATCGCGAGGGCCTCATCAGACAGGAAAAGTGGAAAGCT	133
Qy	135	CAGCACAATCTGGCCAGTGCACCAACAGACAGGAACCTATGTCACTTTCAAGCTCTGGGT	194
Db	134	CAGCACAATCTGGCCAGTGCACCAACAGACAGGAACCTATGTCACTTTCAAGCTCTGGGT	193
Qy	195	TGCGGACACGCATCTCCACACAGGAGTGTCTGCAGAGCTGCCTGCTCGTCATCTTTGGC	254
Db	194	TGCGGACACGCATCTCCACACAGGAGTGTCTGCAGAGCTGCCTGCTCGTCATCTTTGGC	253

JOURNAL  
COMMENT

Unpublished  
On Feb 13, 2001 this sequence version replaced gi:12788604.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC005BF03QPl&cluster=4390.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fullenath.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DC005BF03QPl.  
Location/Qualifiers  
1. .873

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC005YK06"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 165 a 270 c 288 g 149 t

Query Match 59.4%; Score 794.6; DB 9; Length 873;  
Best Local Similarity 99.8%; Pred. No. 7.7e-151;  
Matches 805; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 16 TCCGAGGCAACAGATGGCAGCTGCGGAGCGCTCTCCGGCGCGTTCGTGGGCG 75  
DB 68 TCCGAGGCAACAGATGGCAGCTGCGGA-SCGTCTCCGGCGCGTTCGTGGGCG 126  
QY 76 CGGGCCGCATGGCGGGGCGCATCGCGAGGGCTCTATCAGACAGAGAAAGTGAAGCTC 135  
DB 127 CGGGCCGCATGGCGGGGCGCATCGCGAGGGCTCTATCAGACAGAGAAAGTGAAGCTC 186  
QY 136 AGCACATCTGCGCAGTGACCAACAGACAGGAACCTATGTCACTTTCAAGCTCTGGGT 195  
DB 187 AGCACATCTGCGCAGTGACCAACAGACAGGAACCTATGTCACTTTCAAGCTCTGGGT 246  
QY 196 GCGGACACGACCTCCACACAGGAGTGCTCGACAGCTGCTGCTGCTCATCTTTGCCA 255  
DB 247 GCGGACACGACCTCCACACAGGAGTGCTCGACAGCTGCTGCTGCTCATCTTTGCCA 306  
QY 256 CCAAGCCTCATGTGCTGCCAGTGTCTTGCCAGAGTGGCTCTGCTGCTCACCCTGAAC 315  
DB 307 CCAAGCCTCATGTGCTGCCAGTGTCTTGCCAGAGTGGCTCTGCTGCTCACCCTGAAC 366  
QY 316 ACATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375  
DB 367 ACATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426  
QY 376 CAAACACACGGTGTGCGGGTCTGCGGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 435  
DB 427 CAAACACACGGTGTGCGGGTCTGCGGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 486  
QY 436 TAGTGATGCGCGGGGCGGCCACGTCGCGGAGCAGGACCAAGCTCTGCGAGCTATGCG 495  
DB 487 TAGTGATGCGCGGGGCGGCCACGTCGCGGAGCAGGACCAAGCTCTGCGAGCTATGCG 546  
QY 496 TCGAGGCTGTGGGCGGTGTGAGGAGTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 555  
DB 547 TCGAGGCTGTGGGCGGTGTGAGGAGTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 606  
QY 556 TCAGTGGCAGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615  
DB 607 TCAGTGGCAGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666

QY 616 TCAAGATGGCATGCCAGCAGCCCTGGCCACCGCATCGCTGCCAGACCTGCTGTGGGA 675  
DB 667 TCAAGATGGCATGCCAGCAGCCCTGGCCACCGCATCGCTGCCAGACCTGCTGTGGGA 726  
QY 676 CGGCCAAGATGCTGTGCACGAGGGCCCAACCCAGCCAGCCAGCTGCTGACACGTGTGCA 735  
DB 727 CGGCCAAGATGCTGTGCACGAGGGCCCAACCCAGCCAGCTGCTGACACGTGTGCA 786  
QY 736 CCCCGGGTGGCACCACCATCTATGACTCCACGCCCTGGAGCAGGCGGGCTGCGAGCAG 795  
DB 787 CCCCGGGTGGCACCACCATCTATGACTCCACGCCCTGGAGCAGGCGGGCTGCGAGCAG 846  
QY 796 CCACCATGAGCGCCGTGGAGGTGCCA 822  
DB 847 CCACCATGAGCGCGCTGGAGGTGCCA 873

RESULT 7  
BI837148  
LOCUS

DEFINITION BI837148 873 bp mRNA linear EST 04-OCT-2001  
603089920F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5228956 5',  
mRNA sequence.

ACCESSION BI837148  
VERSION BI837148.1 GI:15948698  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 873)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM11576 row: g column: 05  
High quality sequence stop: 863.

FEATURES  
source

Location/Qualifiers  
1. .873  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5228956"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_120"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-SPORT6; Site:1; NotI; Site:2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dr primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."  
BASE COUNT 154 a 278 c 290 g 151 t

## ORIGIN

Query Match 58.8%; Score 786.4; DB 12; Length 873;  
Best Local Similarity 98.1%; Pred. No. 3.5e-149;  
Matches 859; Conservative 0; Mismatches 11; Indels 6; Gaps 6;  
QY 17 CCGAGGCACAAATGCGAGCTCGGAGCGCTCTCCGCGCGC-GTGGGCTTCGTGGGCG 75  
DB 1 CCGAGGCACAAATGCGAGCTCGGAGCGCTCTCCGCGCGC-GTGGGCTTCGTGGGCG 60  
QY 76 CCGGCGCGCTGGCGGGGCCCATCGCGCAGGGCTCATCAGACAGAGAAAGTGAAGCTC 135





Tel: 81-438-52-3975		Email: genomics@hri.co.jp	
Fax: 81-438-52-3986		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
FEATURES			
source			
1. .794			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/cloned="THYRO1000934"			
/tissue_type="thyroid gland"			
/closure_lib="THYRO1"			
/note="Vector: pME18SFL3"			
BASE COUNT	142 a	248 c	264 g
ORIGIN		133 t	7 others
Query Match	55.1%	Score 737.6;	DB 9; Length 794;
Best Local Similarity	97.7%;	Pred. No. 2.6e-139;	
Matches	765;	Conservative	0; Mismatches 16; Indels 2; Gaps 2;
QY	15	GTCCGAGGCAACAAGATGGCAGCTGCGGAGCCGTCTCCGCGGCGCGTGGCTTCGTGGGC	74
Db	12	GTNCGAGGCAACAAGATGGCAGCTGCGGAGCCGTNTCCGCGGCGCGTGGCTTCGTGGGC	71
QY	75	GCGGGCCGCATGCGCGGGGCCATCGCGAGGGCCTCATCAGACGACGAAAGTGGAAAGCT	134
Db	72	GCGGGCCGCATGCGCGGGGCCATCGCGAGGGCCTCATCAGACGACGAAAGTGGAAAGCT	131
QY	135	CAGCACATACTGGCCAGTGCACCAACAGACAGGAACCTATGTCACTTTCAAGCTCTGGGT	194
Db	132	CAGCACATACTGGCCAGTGCACCAACAGACAGGAACCTATGTCACTTTCAAGCTCTGGGT	191
QY	195	TGCGGACCAACGACTCCAAACGAGGAGGTGCTGCAGAGCTGCCTGCTGTCATCTTTGCC	254
Db	192	TGCCGGACCAACGACTCCAAACGAGGAGGTGCTGCAGAGCTGCCTGCTGTCATCTTTGCC	251
QY	255	ACCAAGCCTCATGTGCTGCCAGCTGTCTGCGAGAGGTGCTCTGTGGTACCACTGAA	314
Db	252	ACCAAGCCTCATGTGCTGCCAGCTGTCTGCGAGAGGTGCTCTGTGGTACCACTGAA	311
QY	315	CACATCTTGGTTCGCTGGCTGCTGGGTGTCTGTGAGCACCTGAGAGAGCTGCTGCC	374
Db	312	CACATCTTGGTTCGCTGGCTGCTGGGTGTCTGTGAGCACCTGAGAGAGCTGCTGCC	371
QY	375	CAAAACACACGGTGCTCGGGGTCTTGCCCAACCTGCCCTGTGTGTCAGGAGGGGC	434
Db	372	CAAAACACACGGGTGCTCGGGGTCTTGCCCAACCTGCCCTGTGTGTCAGGAGGGGC	431
QY	435	ATAGTGTATGGCGGGCGCCACGTGGGAGACAGCGAACAGTCC-TGCAGCATCT	493
Db	432	ATAGTGTATGGCGGGCGCCACGTGGGAGACAGCGAACAGTCC-TGCAGCATCT	491
QY	494	GCTGAGGCGCTGTGGCGGTGTGAGGAGGTGCCTTGAAGCCTACGTGCACATCCACATGG	553
Db	492	GCTGAGGCGCTGTGGCGGTGTGAGGAGGTGCCTTGAAGCCTACGTGCACATCCACATGG	551
QY	554	CCTCAGTGGCAGTGGCGGTTCGTGTGTGTCATTCCTCGAGGCCCTGGCTGGAAGGAGC	613
Db	552	CCTCAGTGGCAGTGGCGGTTCGTGTGTGTCATTCCTCGAGGCCCTGGCTGGAAGGAGC	611
QY	614	CGTCAAGATGGCATGCCACGACGCTTGCCCAACCGCATCGCTGCCACACCTGCTGGG	673
Db	612	CGTCAAGATGGCATGCCACGACGCTTGCCCAACCGCATCGCTGCCACACCTGCTGGG	671
QY	674	GACGCCAAGATGCTGTGTCAGGAGGCCAACCCAGGCCAGCTGCGCTCAGACGCTGTG	733
Db	672	GACGCCAAGATGCTGTGTCAGGAGGCCAACCCAGGCCAGCTGCTGCTCAGACGCTGTG	731
QY	734	CA-CCCGGGTGGCACCAACATCTATGGAATCCACGCCCTTGAGCAGGCGGGCTGCGAG	792

```
Db 239 GGCCTGGCTGAAGAGCGCTCAAGATGGGCTATGGCCAGAGCGCTGGGCCACCGCATCGC 298
Qy 656 TGCCAGAGCCCTGCTGGGACGCCAAGATGCTGTGCACAGAGGGCCAAACACCCAGCCCA 715
Db 299 TGCCAGAGCCCTGCTGGGACGCCAAGATGCTGTGCACAGAGGGCCAAACACCCAGCCCA 358
Qy 716 GCTGCGCTCAGAGCTGTGACCCCGGGTGGCACCACCAATATATGAGTACCTCCAGCCCTGGA 775
Db 359 GCTGCGCTCAGAGCTGTGACCCCGGGTGGCACCACCAATATATGAGTACCTCCAGCCCTGGA 418
Qy 776 GCAGGGGGGGCTGCGAGAGCACCACCATGAGCGCGTGGAGGCTGCCACCTGCGGGGCCAA 835
Db 419 GCAGGGGGGGCTGCGAGAGCACCACCATGAGCGCGTGGAGGCTGCCACCTGCGGGGCCAA 478
Qy 836 GGAGCTCAGCAGAAAGTAGGCTGGGCTCTGGGCATCTCTTCTGCTCTGTGCCCTGCC 895
Db 479 GGAGCTCAGCAGAAAGTAGGCTGGGCTCTGGGCATCTCTTCTGCTCTGTGCCCTGCC 538
Qy 896 TCTCCCTGTGCTCCCTTCCCTGAGGACTCGGCTCCCTCCCTCCCTGATGAGGCTCCT 955
Db 539 TCTCCCTGTGCTCCCTTCCCTGAGGACTCGGCTCCCTCCCTCCCTGATGAGGCTCCT 598
Qy 956 ACTGCTCTCTCCCTTCCCTGAGGACTCGGCTCCCTCCCTCCCTGATGAGGCTCCT 1015
Db 599 ACTGCTCTCTCCCTTCCCTGAGGAAATGAGGGGCGAGGACTTGGAGGTTCAAGA 658
Qy 1016 GCGGGGGGAGCCCGGACAGTGGGACACTCTCCCTCCCTGAGCAGAGGACACCGT 1075
Db 659 GCGGGGGGAGCCCGGACAGTGGGACACTCTCCCTCCCTGAGCAGAGGACACCGC 718
Qy 1076 GGTGGTGG --- CTTCTCCCTTGTCTGCTGAGTGGGACCTCTGCTGCAACATT -GGTCT 1130
Db 719 GGGGGGTGCTCTTGGCCCTTGTCTGAGAGGGAGCCACCTGCTGCCATTTGGTTTG 778
Qy 1131 GAGGGCCCAAGAGATGCTGCTGCTATTTGGCCGCTGTTGGGAGTGGTGGAG 1190
Db 779 AGNGGGCCCAAGAGAGCGGTTTGGGCCATTTGCCCCATGGTGGGAGTTTGTGG 838
Qy 1191 CCATGACAGAACTTACGTTAAAGCAGCGGCTGGCCCAATCCCTGCTGAGGCTGGAG 1250
Db 839 CCATG -ACAGAACTTACGTTAAAGCAGCGGCTGGCCCA -GCCCGGGTGGAACTTGAG 896
Qy 1251 CTTGCTTTGGCTTTTCAAGTGGGCTGCTGAGCTACAGCAGCGGCGGCTCAT 1307
Db 897 CTTGCTTTGGCTTTTCAAGTGGGCTGCTGAGCTACAGCAGCGGCGGCTCAT 953

RESULT 12
BU181305
LOCUS BU181305
DEFINITION AGENCOURT_7935028 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6009767
5', mRNA sequence.
ACCESSION BU181305
VERSION BU181305.1 GI:22695289
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 884)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DPG/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13196 row: 1 column: 24
```

```
FEATURES
source
High quality sequence stop: 428.
Location/Qualifiers
1..884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6009767"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 160 a 281 c 301 g 141 t 1 others
ORIGIN
Query Match 52.7%; Score 704.8; DB 13; Length 884;
Best Local Similarity 96.6%; Pred. No. 1.2e-132;
Matches 762; Conservative 0; Mismatches 23; Indels 4; Gaps 4;
Qy 15 GTCCGAGGACACAAATGGCAGCTGCGGAGCCCTCTCCGCGGCGCTGCTCGTGGC 74
Db 9 GTCCGAGGACACAAATGGCAGCTGCGGAGCCGCTCTCCGCGGCGCTGCTCGTGGC 68
Qy 75 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGGAGCT 134
Db 69 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGGAGCT 128
Qy 135 CAGCACATCTAGGCGGCTGACCAACAGACAGAGAACCTATGTCACTTCAAGCTCGGCT 194
Db 129 CAGCACATCTAGGCGGCTGACCAACAGAGAACCTATGTCACTTCAAGCTCGGCT 188
Qy 195 TGCCGGACACGCACTCAACACAGAGGTGCTGAGAGCTGCTGCTGCTGCTGCTGCT 254
Db 189 TGCCGGACACGCACTCAACACAGAGGTGCTGAGAGCTGCTGCTGCTGCTGCTGCT 248
Qy 255 ACCAAGCCTCATGTGCTGCCAGTGTCTTGGCAGAGTGGCTCTGTGCTGCTGCTGCTGCT 314
Db 249 ACCAAGCCTCATGTGCTGCCAGTGTCTTGGCAGAGTGGCTCTGTGCTGCTGCTGCTGCT 308
Qy 315 CACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374
Db 309 CACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
Qy 375 CCAACACACAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
Db 369 CCAACACACAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
Qy 435 ATAGTGTGCGCGGGCGCCACAGTGGGAGAGCGAGACCAAGCTTCTGCGAGCATCTG 494
Db 429 ATAGTGTGCGCGGGCGCCACAGTGGGAGAGTGGCTGAGAGCTGCTGAGAGCTGCTG 488
Qy 495 CTGAGGCTGCTGCGGGCGGTGTGAGGAGTGGCTGAGAGCTGCTGAGAGCTGCTGAGAG 554
Db 489 CTGAGGCTGCTGCGGGCGGTGTGAGGAGTGGCTGAGAGCTGCTGAGAGCTGCTGAGAG 548
Qy 555 CTCAGTGGCAGTGGCGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
Db 549 CTCAGTGGCAGTGGCGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608
Qy 615 GTCAGATGGGATGCGCCAGAGCTTGGCCCA -CCGATCGCTGCCAGACCTGCTGCTGGG 673
Db 609 GTCAGATGGGATGCGCCAGAGCTTGGCCCA -CCGATCGCTGCCAGACCTGCTGCTGGG 668
Qy 674 GACGGCCAAAGTGTGCTGCTGACAGGGCCCAACACCCAGCCAGCTGGCTGACAGCTGTG 733
Db 669 GACGGCCAAAGTGTGCTGCTGACAGGGCCCAAC -CCAGCCAGCTGGCTGACAGCTGTG 727
Qy 734 CACCCC -GGGTGGCACCACCATCTAT -GGACTTCCAGCCCTGAGAGAGGGCGGCTGCGA 791
Db 728 CACCCC -GGGTGGCACCACCATCTAT -GGACTTCCAGCCCTGAGAGAGGGCGGCTGCGA 787
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 103 Seconds  
(without alignments)  
5733.695 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcagctctccga.....ggggcagcagcatatggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTOUT\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201.8	15.1	1742	3	US-09-099-676-2
2	201.8	15.1	1742	3	US-09-365-910-2
C 3	91.8	6.9	879	4	US-09-252-991A-359
C 4	91.8	6.9	1164	4	US-09-252-991A-385
5	91.8	6.9	1419	4	US-09-252-991A-440
C 6	88	6.5	42988	4	US-08-311-731A-128
7	75.8	5.7	28171	4	US-08-961-527-22
8	73.4	5.5	4403765	3	US-09-103-840A-2
9	73.4	5.5	4411529	3	US-09-103-840A-1
10	66.6	5.0	1147	1	US-08-665-716-1
11	58.6	4.4	516	4	US-09-252-991A-4140
C 12	51.4	3.8	1116	4	US-09-252-991A-2320
13	51.4	3.8	2082	4	US-09-252-991A-2596
14	51.4	3.8	2388	4	US-09-252-991A-2514
15	48.8	3.6	633	4	US-09-252-991A-488
C 16	47.6	3.6	1686	4	US-09-252-991A-4056
17	46.4	3.5	1230	4	US-09-252-991A-16191
C 18	46.4	3.5	1812	4	US-09-252-991A-16295
19	46.4	3.5	2019	4	US-09-252-991A-15983
20	44.6	3.3	1746	4	US-09-252-991A-4395
C 21	44.6	3.3	2571	4	US-09-252-991A-4351
22	44.6	3.3	2997	4	US-09-252-991A-4487
C 23	44	3.3	1645	4	US-09-620-312D-807
24	42.8	3.2	466	3	US-09-154-083-17
25	42.2	3.2	3978	4	US-09-266-965-19
26	42.2	3.2	7218	1	US-08-232-463-14
27	42.2	3.2	12249	4	US-09-266-965-74

28	42.2	3.2	18331	4	US-09-266-965-96
C 29	42	3.1	1155	4	US-09-252-991A-12035
C 30	42	3.1	1185	4	US-09-252-991A-12007
31	42	3.1	1407	4	US-09-252-991A-11979
C 32	41.6	3.1	1365	4	US-09-252-991A-1996
C 33	41.6	3.1	1389	4	US-09-252-991A-1733
C 34	41.2	3.1	870	4	US-09-252-991A-2385
C 35	41.2	3.1	1326	4	US-09-252-991A-2299
C 36	41.2	3.1	1926	4	US-09-249-585A-2
C 37	41.2	3.1	1926	4	US-09-410-399-3
C 38	41.2	3.1	2580	3	US-09-050-863-2
C 39	41.2	3.1	2580	4	US-09-359-081-2
40	41.2	3.1	5452	2	US-09-130-114-1
41	41.2	3.1	8705	4	US-09-647-344A-14
C 42	41.2	3.1	9600	3	US-08-910-647-1
C 43	41.2	3.1	9600	4	US-09-620-925-1
C 44	41.2	3.1	10596	1	US-07-884-811-15
C 45	41.2	3.1	10596	1	US-07-885-971-15

## ALIGNMENTS

RESULT 1  
US-09-099-676-2  
; Sequence 2, Application US/09099676  
; Patent No. 6100075  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE  
; TITLE OF INVENTION: HOMOLOG  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/099,676  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrone, Michael C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0532 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-855-0572  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1742 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNON01  
; CLONE: 2278458  
US-09-099-676-2

Query Match 15.1%; Score 201.8; DB 3; Length 1742;





Db 319 CGGCGCGGCGTATTTCTTCTGCTGATGAGGCGCATGACCGAGCGCGGAGAGCTCG 260  
QY 625 GCATGCCAGAGAGCTGGCCACCGCATCGTGCACAGACCCCTGCTGGGACGCGCAAGA 684  
Db 259 GCCTGTCCGGGAGACAGCCTCGCGCTGACCCCTGCAGACCGCCCTCGGCGCGGCGCAGA 200  
QY 685 TGCTGCTGCAGAGGGGCAACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTG 744  
Db 199 TGGCGCTGTCCAGGAGGTGAGCGCGCGCAACTGCGCGCGCGGTGACTTCGCCCAACG 140  
QY 745 GCACCAACCATATATGAGTCCACGCCCTGAGAGGCGGCGTGCAGAGAGCCACCATGA 804  
Db 139 GCACCAAGGCGGCGCATCAATCTTCCAGGCGCAAGCGCTTCGAAGCACTGGTGAGC 80  
QY 805 GCGCGGTGAGGCTGCGCACTCGCGGCGCAAGGAGTTCAGCAGAAAG 851  
Db 79 AGGCCCTGAAGCGCGCGCAGCAGCGCTCCGCGAGCTGGCGCAACAG 33

## RESULT 5

US-09-252-991A-440  
; Sequence 440, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-07-88  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 440  
; LENGTH: 1419  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-440

Query Match  
Best Local Similarity 6.9%; Score 91.8; DB 4; Length 1419;  
Matches 307; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

QY 208 ACTCCAAACGAGGTGCTGCAGAGTGCCTGCTGCTATCTTGCACCAAGCCTCATG 267  
Db 761 AGTCCAAACGAGGCGGTGGCGGACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 820  
QY 268 TGTGCGCAGTGTCTGCGCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327  
Db 821 CCATGAAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
QY 328 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387  
Db 881 CCATGCGCGGATCCCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940  
QY 388 TGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447  
Db 941 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000  
QY 448 GGGGCGCGCAGTGGGAGCAGCAGCAAGCTCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 507  
Db 1001 CCACGCGGAGTGCAGCGCGCGCAACGCGGAGCGGCGGCGGAGTGTGCTGCTGCTGCTGCT 1060  
QY 508 GCGGCTGTGAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564  
Db 1061 GCATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120  
QY 565 GTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624  
Db 1121 CGGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180

QY 625 GCATGCCAGAGCCTGGCCACCGCATGCTGCCAGACCTGCTGGGAGCGGCAAGA 684  
Db 1181 GCCTGTCCGGGAGACAGCCTCGCGCTGACCCCTGCAGACCGCCCTCGGCGGCGCAGA 1240  
QY 685 TGCTGCTGCAGAGGGGCAACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTG 744  
Db 1241 TGGCGCTGTCCAGGAGGTGAGCGCGCGCAACTGCGCGCGCGGTGACTTCGCCCAACG 1300  
QY 745 GCACCAACCATATATGAGTCCACGCCCTGAGAGGCGGCGTGCAGAGAGCCACCATGA 804  
Db 1301 GCACCAAGGCGGCGCATCAATCTTCCAGGCGCAAGCGCTTCGAAGCACTGGTGAGC 1360  
QY 805 GCGCGCTGAGGCTGCCACCTGCGGCGCAAGAGCTCAGCAGAAAG 851  
Db 1361 AGGCCCTGAAGCGCGCGCAGCAGCGCTCCGCGAGCTGGCGCAACAG 1407

## RESULT 6

US-08-311-731A-128/c  
; Sequence 128, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,731A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: C0044/7125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/720-3500  
; TELEFAX: 617/720-2441  
; INFORMATION FOR SEQ ID NO: 128:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42988 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: MYCOBACTERIUM LEPRAE  
US-08-311-731A-128

Query Match  
Best Local Similarity 50.7%; Pred. No. 2.3e-11;  
Matches 211; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 279 GTCTGCGCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 338  
Db 27503 GTTCAGGCGCGCAGTGCAGCAATGACAGCGCTGAGCAAGTGTAGTCACCGTGGCTGCT 27544



QY	344	GTCTCTGACGACCCCTGGAGGAGCTGCTGCCCCCCAAACACACAGGGGTGCTGGGGTCTTTGCC	403
Db	591763	CACGATCGCGTATTTTCGAATCCAAGTACCGCCGGGACGCCAGTGGTGGTCGGATGCC	591822
QY	404	CAACCTGGCCCTGTGTGCTCCAGGAGGGCCATAGTATGGCGGGGGCGCCACGTGGG	463
Db	591823	GAACGCGGCGCATTTGTGTGGAGCGGGGGTTACAGCGCTGCTGCCAAAGCGCCTTTGTCCAC	591882
QY	464	GAGCAGCGAGACCAAGCTCCTTGCAGCATCTCTCTGGAGGSCCTGTGGCGGGTGTGAGAGGT	523
Db	591883	CCGGCACAAGCTTGAGGAGGTCTCGGCCCTTTTCGACGCGGTGGCGGGGTGCTGACCGT	591942
QY	524	GCGTGAAGCCTACGTGCGACATCCACTGGCCCTCAGTGGCAGTGGCGTGGCCCTTCGTG	583
Db	591943	TCCGGAATCGCATTTGGACGCGGTGACCGCGGTCTCCGGGTCCGGGTCCGGCCTATTCTT	592002
QY	584	TGATTTCTCGAGGCCCTTGGCTGAAGGAGCGGTCAAGATGGGCATGCCACGACCGCTGCG	643
Db	592003	TCTGTGTCGAGGCGCCTGGTGGATGCCGGAGTCGGGTGGGCTTGGCCGTGAGCCGTGAGTGGC	592062
QY	644	CCACGCGATCGTCCCGACAGCCCTGTCTGGGACGGCCAAAGATGCTGCTGCA	694
Db	592063	CACCGATCTCGCGCGCAGCAATGGTGGCTACGCGGCATGCTGTGGA	592113

```

RESULT 9
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	5.5%; Score 73.4; DB 3; Length 441529;
Best Local Similarity	48.7%; Pred. No. 2.2e-07;
Matches 200; Conservative 0; Mismatches 211; Indels 0; Gaps 0;	
QY	284 GCGCAGGTGGCTCTGTGGTTCACCATCTGAACACATCTTGGTGTCCGTGGCTGCTGGGGT 343
Db	
	590341 GACTCGCGCGCGCGAAGAACACACAGTGTGTGACGAGGTTCGTACCGCTGGTACGCGCAT 590400
QY	344 GTCTCTGAGCACCCCTGGAGGAGCTGTGCCCCCAACACACGSGGTCTCGGGCTTGGCC 403
Db	
	590401 CACGATCCGCTATTTCGAATCCAGCTACCGGCTGGGACGCCAGTGGTCGTGGCATGGC 590460
QY	404 CAACCTGCCCTGTGTGTCTCCAGGAAGGGCCATAGTATGGCGCGGGCGCCAGCAGTGGG 463
Db	
	590461 GAACCGCGCGCATTTGTTGGGAGCGGGGTTACAGCGCTGGCCAAAGGCCGTTTGTTCAC 590520
QY	464 GAGCAGCGAGACCAAGCTCTCTGCAGCATCTGTCTGGAGGCGCTGTGGCGCGTGTGAGAGGT 523
Db	
	590521 CCGCGCAACAGCTTGTAGGAGAGTCTGGCGCTTGTTCGACGCGGTGGCGCGGTGCTACCGT 590580
QY	524 GCCTGAAGCCCTACCTGTCGACATCCACACTGSCCCTCAGTGGCGAGTGGCGGTTCGCTTGTG 583
Db	
	590581 TCCGGAATCGAGTTGTGACGCGGTGACCGCGTCTCCGCTCTGGGTCCGGCCCTATTCTT 590640
QY	584 TGCATTTCCGAGGCCCTGGCTGAAGAGGCCGTCAAGATGGGCATGCCACAGCCTGGC 643

```

Db      590641  TCTGCTGGTGGAGGCCCTGGTGGATGCCGGAGTCGGGGTGGGCTTGAGCCGTCAGGTGGC 590700
QY      644    CCACCGCATCGTGGCCAGACCCCTGCTGGGACGCGCCCAAGATGCTGCTGCA 694
Db      590701  CACCGATCTCGCGCGCGACACAATGGCTGGCTCAGCGCGCATGCTGCTGGA 590751

RESULT 10
US-08-665-716-1
; Sequence 1, Application US/08665716
; Patent No. 5789222
; GENERAL INFORMATION:
; APPLICANT: KELLY, ROSEMARIE
; APPLICANT: REGISTER, ELIZABETH A
; APPLICANT: MASUREKAR, PRAKASH S
; TITLE OF INVENTION: P5C REDUCTASE GENE FROM ZALERION
; TITLE OF INVENTION: ARBORICOLA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: US
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,716
; FILING DATE: 23-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KORSEN, ELLIOTT
; REGISTRATION NUMBER: 32,705
; REFERENCE/DOCKET NUMBER: 19453PV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-5493
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..960
US-08-665-716-1

```

Query Match	5.0%;	Score 66.6;	DB 1;	Length 1147;
Best Local Similarity	52.3%;	Pred. 9.4e-07;		
Matches 172;	Conservative 0;	Mismatches 154;	Indels 3;	Gaps 1
QY	520	AGTGCCTGAAGCCTACGTGCACATCCACACTCCACACTGCCTCAGTGCACCTGCCTGGCTTCGG	579	
DB	651	AGTCCCCGTGCTACATGTGATGCCACACCGCGCTGTGTGCATGTGCCAGCGCTTTT	710	
QY	580	TGTGTGCATTCTCCGAGGCGCCTGGCTCAAGGAGCGCTCAAGATGGCCTACCCACGACGCC	639	
DB	711	TTGCGCTCATACTGGAGGCTCGATTATGATGGGCTGTGGCGATGGGGTTGCCAAGACGAC	770	
QY	640	TGGCCCCCGCATCGCTGCCACGACCTGCTGGGGAGCGGCCAAGATGCTGTCTGCACGAGG	699	
DB	771	AGGCTCAGAGAAATGGCGCGCAGACTATGAAGAGGGCTGCTGGATTGTTATTGTCTGGAG	830	
QY	700	GCACACCCAGCCGACGCTGCGCTCAGACCTGTGCACCCCGGGTGGCACACCATCTATG	759	



Db 831 ---AGCATCGGGCTTCTGAAAGATAAAAGTACTAGCCGGGTGGGTGCACGATAGGTG 887  
QY 760 GACTCCAGCCCTGGAGCAGGGCGGCTGCAGCAGCCACCATGAGCGCGTGGAGGCTG 819  
Db 888 GGCTGATGGTCTGGAAGAGGAGGGGTGAGAGGGACCGTGGCTAGAGCAGTTAGGGAAG 947  
QY 820 CCACCTCCGGGCGCAAGGAGCTCAGCAGA 848  
Db 948 CTACTGTGGTGGCAGTCAGCTTGGTAAA 976

RESULT 11  
US-09-252-991A-4140  
; Sequence 4140, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4140  
; LENGTH: 516  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4140

Query Match 4.4%; Score 58.6; DB 4; Length 516;  
Best Local Similarity 46.9%; Pred. No. 5.7e-05;  
Matches 220; Conservative 0; Mismatches 244; Indels 5; Gaps 1;  
QY 388 TGCTGGCGGTCTTGCCCAACCTGCGCTGTGTGTCAGGAGGCGGCATAGTCATGCGC 447  
Db 36 TGGCGCTGGGGGGCGCGCTCCAGCTGTGGCCACCGCGCGGGGACGCTTCGTAG 95  
QY 448 GGGGCCCGCAGTGGGGAGCAGCAGACCAAGCTCTCTGCAGCATCTGCTGGAGGCGCTGTG 507  
Db 96 GTCGGCGCCATGGAGCAGCGCGCGAGGGTCAGGGCGATCAGCAGCGCGATACGCCGAAT 155  
QY 508 GCGCGGTGAGGAGTCCCTGAAGCCTACGTGCACATCCACATGCGCTCAGTGGCAGTG 567  
Db 156 GCAGGTTTGGCGATCTCGTTCACTCCCGCGCCGAAGCGGCTGTCGATGGGTCCGGCT 215  
QY 568 GCGTGGCTTCGTGTGCATTTCTCC----GAGGCGCTGGCTGAAGGAGCGCGTCAAGAT 622  
Db 216 TGCTGGCAGCAGCATTTCTTCTGTGTGATGAGGCCATACCCAGCGCGCGGAGAGCT 275  
QY 623 GGGCATGCCAGCAGCCTGGCCACCCATCGCTGCGCAGACCCCTGCTGGGGACGGGCCAA 682  
Db 276 CGGCTGTCCCGGAGACAGCCTCGCGCCTGACCTCAGACACCGCCCTCGCGCGGCGCA 335  
QY 583 GATGCTGCTGCAGAGGGCCCAACACCCAGCCCGAGCTGCGCTCAGAGTGTGACCCGGG 742  
Db 336 GATGGCGCTGTCCAGCAGGTCGAGCGCGCGCACTGCGCGCGGGTGACTTTCGCCCAA 395  
QY 743 TGGCACACCATCTATGAGCTCCACGCCCTGGAGCAGGGCGGCTGCGAGCAGCCACCAT 802  
Db 396 CGGCACACCGAGCGCGGATCAATCTTCCAGGCCAACGGCTTGAAGCATCTGGTGA 455  
QY 803 GAGCGCGTGGAGGCTGCCACCTGCGCGGCCCAAGAGGCTCAGCAGAAAG 851  
Db 456 GCAGGCCCTGAACGCCCGCCAGCGCGCTCCGCCGAGCTGGCCGACAG 504

RESULT 12  
US-09-252-991A-2320/c

; Sequence 2320, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2320  
; LENGTH: 1116  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-2320

Query Match 3.8%; Score 51.4; DB 4; Length 1116;  
Best Local Similarity 46.1%; Pred. No. 0.0035;  
Matches 172; Conservative 0; Mismatches 201; Indels 0; Gaps 0;  
QY 186 GCTCTGGGTTCGGGACCAACGCACTCCAAACAGGAGGTGCTGCAGAGCTGCTGCTGCTC 245  
Db 822 GCCCTGGCCTGCTGTTGCTGTCGTGGCGCGGGAACCTCGCATCTCGCTCCGCGCC 763  
QY 246 ATCTTTGCCACCAAGCCTCATGTGTCGCCAGCTGTCCTGGCAGAGTGGCTCCTGTGGTC 305  
Db 762 CTCGCCGAAGTCCGCGCATCGATTCGGGTGGCATCGCTGCACCCCTGCTGGGAGTC 703  
QY 306 ACCACTGAACACATCTTGGTGTCCGTGGCTGCTGGGTGCTCTCTGAGCACCTCGAGGAG 365  
Db 702 ATCATCGGGGTGATTTGCTGACCGCATCGCCACACCTTCGCCAGCTACGCTGCTCGAC 643  
QY 366 CTGCTGCCCCCAACACACACGCGGTGCTGCGGGTCTTTGCCCAACCTGCCCTGTGTGCCAG 425  
Db 642 CTGGGACGGGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583  
QY 426 GAAGGGCCATAGTATGTCGGCGGGCGGCCCGCCAGCTGGGGAGCAGGACCAAGCTCCTG 485  
Db 582 GGAATGGGATCCCGACCATCCCAACTACATCATCACAGCTCCATCGCCGCGCGCG 523  
QY 486 CAGCATCTCTCGAGGSCCTGTGGCGGTGTGAGGAGGTGCTGAAGCCTGACGCTCGACATC 545  
Db 522 CTGCTGGAAGTGGGGTGGCGCTGATCTCGACATGTTCTTCTTCTTCTTCTTCTTCTT 463  
QY 546 CACACTGGCCTCA 558  
Db 462 CTCGCGGACCTCA 450

RESULT 13  
US-09-252-991A-2596  
; Sequence 2596, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2596  
; LENGTH: 2082  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa





BASE COUNT	163 a	311 c	306 g	152 t
ORIGIN				
Query Match	52.1%; Score 697.4; DB 13; Length 932;			
Best Local Similarity	97.0%; Pred. No. 3.9e-131;			
Matches 732; Conservative	0; Mismatches 21; Indels 2; Gaps 2;			
Qy	41	GGAGCGCTCTCCGCGCGGTGGGCTTCGTGGCGCGGCCGCGCATGGCGGGGCCCATGCG	100	
Db	1	GGAGCGCTCTCCGCGCGGTGGGCTTCGTGGCGCGGCCGCGCATGGCGGGGCCCATGCG	60	
Qy	101	GCAGGGCCTCATCAGACGAGGAAAAGTGAAGCTCAGACATACCTGAGTGGCCAGTGCACCAAC	160	
Db	61	GCAGGGCCTCATCAGACGAGGAAAAGTGAAGCTCAGACATACCTGAGTGGCCAGTGCACCAAC	120	
Qy	161	AGACAGGAACCTATGTCACTTTCAAGCTCTGGTTCGCGGACACAGCACTTCAACACGAGA	220	
Db	121	AGACAGGAACCTATGTCACTTTCAAGCTCTGGTTCGCGGACACAGCACTTCAACACGAGA	180	
Qy	221	GGTGTGCAGAGCTGCTGCTCATCTTTTGCACCAAGCCTCATGTGCTGCCAGCTGT	280	
Db	181	GSTGCTGCAGAGCTGCTGCTCATCTTTTGCACCAAGCCTCATGTGCTGCCAGCTGT	240	
Qy	281	CCTGGCAGAGTGGCTCCTGTGTGCTCACCCTGAAACACATCTTGGTCCGTGGCTGCTGG	340	
Db	241	CTTGGCAGAGTGGCTCCTGTGTGCTCACCCTGAAACACATCTTGGTCCGTGGCTGCTGG	300	
Qy	341	GSTGCTCTGAGCACCTCGAGAGCTGCTGCCGCCCAACACACATCTTGGTCCGTGGCTGCT	400	
Db	301	GATGCTCTGAGCACCTCGAGAGCTGCTGCCGCCCAACACACATCTTGGTCCGTGGCTGCT	360	
Qy	401	GCCCACTGCTGCTGTGTGTCAGGAAGGGCCATAGTGTGGCGGGGCCGACAGT	460	
Db	361	GCCCACTGCTGCTGTGTGTCAGGAAGGGCCATAGTGTGGCGGGGCCGACAGT	420	
Qy	461	GGGAGCAGCAGACCAAGCTCTGCAGCATCTGCTGGAGGCTGCTGGCGGTGTGAGA	520	
Db	421	GGGAGCAGCAGACCAAGCTCTGCAGCATCTGCTGGAGGCTGCTGGCGGTGTGAGA	480	
Qy	521	GSTGCTGAGGCTAGCTGCATCCACACTGSCCTCAGTGGCAGTGGCGTGGCTTGGT	580	
Db	481	GSTGCTGAGGCTAGCTGCATCCACACTGSCCTCAGTGGCAGTGGCGTGGCTTGGT	540	
Qy	581	GTGTGATCTCCGAGGCGCTGGCTGAAGAGCGCTCAAGATGGCGATGCCAGAGCT	640	
Db	541	GTGTGATCTCCGAGGCGCTGGCTGAAGAGCGCTCAAGATGGCGATGCCAGAGCT	600	
Qy	641	GGCCACCGCATCGCTGCCAGACCTGCT-GGGAGCGCCAGATGCTGTGTCACGAGG	699	
Db	601	GGCCACCGCATCGCTGCCAGACCTGCTGGGAGCGCCAGATGCTGTGTCACGAGG	660	
Qy	700	GCCAACACCCAGCCAGCTGCGCTCAGACGTGTGCACCC-CGGTGGCACCCACCATCAT	758	
Db	661	GCCAACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGGGGACCATCAT	720	
Qy	759	GGACTCCAGCCCTGAGCAGGGCGGGCTGCCAGC	793	
Db	721	GGACTCCCGCCCTGGGAACACGAGGGGGGGCTGC	755	
RESULT 15				
BI253559				
LOCUS	602973485F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112692 5',			
DEFINITION	mRNA sequence.			
ACCESSION	BI253559			
VERSION	BI253559.1 GI:14805103			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 734)			

AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L14M11273 row: j column: 21  
High quality sequence stop: 734.

FEATURES

source

1. 734  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5112692"  
/tissue\_type="cervical carcinoma cell line"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_12"  
/note="Organ: cervix; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.4 kb. Library prepared by Life  
Technologies."

BASE COUNT 127 a 228 c 253 g 126 t

Query Match 51.4%; Score 688.2; DB 12; Length 734;

Best Local Similarity 98.6%; Pred. No. 2.6e-129;

Matches 715; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

Qy 36 GCTCGGAGCGCTCCGCGCGGTGGCTTCGTGGCGCGGCCGCGCATGGCGGGGCC 95

Db 11 GGTGCGAGCGCTCCGCGCGGTGGGGTCTGGCGCGGCCGCGCATGGCGGGGCC 70

Qy 96 ATCGCGAGGGCCCTCATCAGACGAGGAAAAGTGAAGCTCAGCACATCTGGCGACGTGCA 155

Db 71 ATCGCGAGGGCCCTCATCAGACGAGGAAAAGTGAAGCTCAGCACATCTGGCGACGTGCA 130

Qy 156 CCAACACAGGAACTATGTCACTTTCAAGCTCTGGGTTGCCGGACCGACGCTCCAC 215

Db 131 CCAACACAGGAACTATGTCACTTTCAAGCTCTGGGTTGCCGGACCGACGCTCCAC 190

Qy 216 CAGGAGTGTGTGACAGCTGCTGCTCATCTTTGCCACCAAGCTCATGTGCTGCCA 275

Db 191 CAGGAGTGTGTGACAGCTGCTGCTCATCTTTGCCACCAAGCTCATGTGCTGCCA 250

Qy 276 GCTGCTCGCAGAGTGGCTTCCTGTTGGTCAACACTGAACACATCTTGTGCTCCGTGGCT 335

Db 251 GCTGCTCGCAGAGTGGCTTCCTGTTGGTCAACACTGAACACATCTTGTGCTCCGTGGCT 310

Qy 336 GCTGGGTGTCTGTGACACCGCTGGAGAGCTGCTGCCCGCCCAACACACAGGGTGTGCGG 395

Db 311 GCTGGGTGTCTGTGACACCGCTGGAGAGCTGCTGCCCGCCCAACACACAGGGTGTGCGG 370

Qy 396 GTCCTGCCCAACCTGCCCTGTGTGTCAGAGGGGCCATAGTGTGCGCGCGGGCGCG 455

Db 371 GTCCTGCCCAACCTGCCCTGTGTGTCAGAGGGGCCATAGTGTGCGCGCGGGCGCG 430

Qy 456 CACGTGGGAGCAGCAGAACAGCTCCTGCAGCATCTGCTGGAGGCGCTGTGGCGGTGT 515

Db 431 CACGTGGGAGCAGCAGAACAGCTCCTGCAGCATCTGCTGGAGGCGCTGTGGCGGTGT 490

Qy 516 GAGGAGTGGCTGAAGCCTAGCTGCAGATCCACACTGGCTCAGTGGCAGTGGCGGCC 575

Db 491 GAGGAGTGGCTGAAGCCTAGCTGCAGATCCACACTGGCTCAGTGGCAGTGGCGGCC 550

Qy 576 TTCGTGTGCATTCTCCGAGGCGCTGGCTGAAGGAGCGCTCAAGATGGGCATGCCAGC 635

Db 551 TTCGTGTGCATTCTCCGAGGCGCTGGCTGAAGGAGCGCTCAAGATGGGCATGCCAGC 610

QY	636	AGCCTGGCCACCGCATCGCTGCCAGACCCCTGCTGGGACGGCCAAAGATGCTGTCAC	695
Db	611	AGCCTGGCCACCGCATCGCTGCCAGACCCCTGCTGGGACGG-CAAGATGCTGTCAC	669
QY	696	G-AGGCCAACAACCCAGCCAGCTGGCTCAGACCTGTGCACCCCGGTGGACACCAT	754
Db	670	GAAGGCCAAAACCCAGCCAGCTGGCTCAGACCTGTGCACCCCGGTGGACACCAT	729
QY	755	CTATG	759
Db	730	CTATG	734

Search completed: August 8, 2003, 17:47:43  
Job time : 2686 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 4867 Seconds  
(without alignments)  
11246.575 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcgcagctgtccga.....ggggcagcagcatatgggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.ph.\*
- 7: gb.pl.\*
- 8: gb.pr.\*
- 9: gb.ro.\*
- 10: gb.sts.\*
- 11: gb.sy.\*
- 12: gb.un.\*
- 13: gb.vi.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_mam.\*
- 37: em.htg\_vrt.\*
- 38: em.sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1307	97.7	2331	6	BD160280	BD160280 Primer fo
2	1307	97.7	2331	9	AK023914	AK023914 Homo sapi
c	1291.8	96.5	2615	6	BD156362	BD156362 Primer fo
c	1291.8	96.5	2615	9	AK001500	AK001500 Homo sapi
5	1200.4	89.7	2400	9	AL833857	AL833857 Homo sapi
6	1127.8	84.3	1178	9	BC007993	BC007993 Homo sapi
7	737.6	55.1	794	6	BD150492	BD150492 Primer fo
c	653.2	48.8	101777	9	AC105210	AC105210 Homo sapi
c	653.2	48.8	184959	2	AC087823	AC087823 Homo sapi
c	653.2	48.8	174906	9	AC067930	AC067930 Homo sapi
11	599.4	44.8	207636	2	AC019122	AC019122 Homo sapi
12	567.6	42.4	1276	10	BC026536	BC026536 Mus muscu
13	486.6	36.4	563	6	BD148256	BD148256 Primer fo
c	338.4	25.3	197243	2	AC120593	AC120593 Rattus no
15	226.4	16.9	1848	9	BC001504	BC001504 Homo sapi
16	224.8	16.8	1792	6	AX409754	AX409754 Sequence
17	224.8	16.8	1792	9	HUMP5CR	M77836 Human pyro
18	213.8	16.0	1769	9	AF218000	AF218000 Homo sapi
19	205.4	15.4	1757	9	BC022244	BC022244 Homo sapi
20	201.8	15.1	1676	9	BC014868	BC014868 Homo sapi
21	201.8	15.1	1708	9	BC020553	BC020553 Homo sapi
22	191.6	14.3	1590	10	BC006882	BC006882 Mus muscu
23	173.4	13.0	1820	10	BC006727	BC006727 Mus muscu
24	167	12.5	999	9	AF151351	AF151351 Homo sapi
25	156.4	11.7	299800	1	AP005040	AP005040 Streptomy
c	151.6	11.3	293050	1	SCO939116	AL939116 Streptomy
27	148.8	11.1	3149	6	BD156370	BD156370 Primer fo
28	148.8	11.1	3149	9	AK001508	AK001508 Homo sapi
c	147	11.0	148283	2	AC116520	AC116520 Mus muscu
c	143.4	10.7	10490	1	AE013306	AE013306 Methanosa
31	143.2	10.7	178292	2	AC120830	AC120830 Rattus no
32	143.2	10.7	245651	2	AC126537	AC126537 Rattus no
33	141.4	10.6	999	3	AY071558	AY071558 Drosophil
34	138.4	10.3	822	1	AY142817	AY142817 Heliohaci
35	138.4	10.3	5139	1	AF305580	AF305580 Methanosa
36	138.4	10.3	10357	1	AE011122	AE011122 Methanosa
c	127.6	9.5	28128	2	AC015301	AC015301 Drosophil
38	127.6	9.5	171612	3	AC009461	AC009461 Drosophil
39	127.6	9.5	175440	3	AC007804	AC007804 Drosophil
40	127.6	9.5	253101	3	AE003718	AE003718 Drosophil
41	126	9.4	10029	1	AE012787	AE012787 Chlorobiu
42	123	9.2	2678	1	TTTHPROC	D25413 Thermus the
43	118.2	8.8	309350	1	AP005215	AP005215 Corynebac
44	116.6	8.7	1089	3	AF098020	AF098020 Drosophil
45	111.8	8.4	4290	3	AF170829	AF170829 Drosophil

ALIGNMENTS

RESULT 1

BD160280

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BD160280 2331 bp DNA linear PAT 17-JAN-2003  
Primer for synthesizing full-length cDNA and use thereof.

BD160280

BD160280.1 GI:27866038

JP 2002191363-A/15123

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Ota.T., Isogai.T., Nishikawa.T., Saito.K., Yamamoto.J.,

Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.

Primer for synthesizing full-length cDNA and use thereof

JOURNAL	Patent: JP 2002191363-A 15123 09-JUL-2002;
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/15123 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OHSUKI PC
FEATURES	C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC 10, PC C12P21/02,C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH key Location/Qualifiers FT CDS (27).(848).
source	Location/Qualifiers 1..2331 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
BASE COUNT	398 a 753 c 748 g 432 t
ORIGIN	
Query Match	97.7%; Score 1307; DB 6; Length 2331;
Best local Similarity	99.2%; Pred. No. 1e-224;
Matches 1313; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
QY	15 GTCCGAGGCAACAAGATGGCAGCTGCGGAGCGCTCTCCGGCGCGGTGGGCTTCGTGGGC 74
Db	12 GTCCGAGGCAACAAGATGGCAGCTGCGGAGCGCTCTCCGGCGCGGTGGGCTTCGTGGGC 71
QY	75 GCGGGCGCATGCGCGGGGCCATTCGCGCAGGGCGCTCATCAGAGCAGGAAAAGTGAAGCT 134
Db	72 GCGGGCGCATGCGCGGGGCCATTCGCGCAGGGCGCTCATCAGAGCAGGAAAAGTGAAGCT 131
QY	135 CACACATACATGCGCAGTGCACCAACAGACAGAACTATGTACATTTCAAGCTCTGGGT 194
Db	132 CACACATACATGCGCAGTGCACCAACAGACAGAACTATGTACATTTCAAGCTCTGGGT 191
QY	195 TGCCGGACACGACCTCCACACAGGAGTGCCTGCAGAGCTGCCTGCTGTCATCTTTGCC 254
Db	192 TGCCGGACACGACCTCCACACAGGAGTGCCTGCAGAGCTGCCTGCTGTCATCTTTGCC 251
QY	255 ACCAAGCCTCATGTGTCGACGTGCTCTGGCAGAGTGGCTCTGTGTCCAGGAGTGGCTCACCCTGAA 314
Db	252 ACCAAGCCTCATGTGTCGACGTGCTCTGGCAGAGTGGCTCTGTGTCCAGGAGTGGCTCACCCTGAA 311
QY	315 CACATCTTGGTGTCCGTGGCTGCTGGGGTGTCTCTGAGCACCTGGAGAGCTGTGCCC 374
Db	312 CACATCTTGGTGTCCGTGGCTGCTGGGGTGTCTCTGAGCACCTGGAGAGCTGTGCCC 371
QY	375 CCAACACACGGGTGCTGGGGTCTTGGCCACCTGCTGTGTGTGGTCCAGGAGGGCC 434
Db	372 CCAACACACGGGTGCTGGGGTCTTGGCCACCTGCTGTGTGTGGTCCAGGAGGGCC 431
QY	435 ATAGTGTATGCGCGGGCGCCACCTGCGGAGCAGCAGAACCTCCTCGAGCATGTG 494
Db	432 ATAGTGTATGCGCGGGCGCCACCTGCGGAGCAGCAGAACCTCCTCGAGCATGTG 491
QY	495 CTGGAGGCGCTGTGGCGGTGTGAGGAGTGTGCTGAAGCTACGTGACATCCACACTGGC 554
Db	492 CTGGAGGCGCTGTGGCGGTGTGAGGAGTGTGCTGAAGCTACGTGACATCCACACTGGC 551
QY	555 CTCAGTGGCAGTGGCGGTGCTTGTGTGTGCAATTCCTCCAGGCGCTGGCTGAAGAGGCC 614
Db	552 CTCAGTGGCAGTGGCGGTGCTTGTGTGTGCAATTCCTCCAGGCGCTGGCTGAAGAGGCC 611
QY	615 GTCAAGATGGGCATGCCAGCAGCTGGCCACCGCATCGTGGCCAGACCTCTGCTGGG 674
Db	612 GTCAAGATGGGCATGCCAGCAGCTGGCCACCGCATCGTGGCCAGACCTCTGCTGGG 671
QY	675 ACGGCCAAGATGCTGCTGCACGAGGGCCAAACACCCAGCCAGCTGCGCTCAGACGTGTC 734
Db	672 ACGGCCAAGATGCTGCTGCACGAGGGCCAAACACCCAGCCAGCTGCGCTCAGACGTGTC 731
QY	735 ACCCGGGTGGCACCACCATCTATGACTTCCACGCCCTGGAGCAGGGCGGCTCGGAGCA 794
Db	732 ACCCGGGTGGCACCACCATCTATGACTTCCACGCCCTGGAGCAGGGCGGCTCGGAGCA 791
QY	795 GCCACCATGAGCGCGGTGGAGGCTGCCACCTGCCGGGCCAAAGAGCTCAGCAGAAAGTAG 854
Db	792 GCCACCATGAGCGCGGTGGAGGCTGCCACCTGCCGGGCCAAAGAGCTCAGCAGAAAGTAG 851
QY	855 GCTGGGCTCTGCGCATCTTCCCTGCTGTGCCCTGTGCCCTGCTCCTCCTGCTGCTCCCTTG 914
Db	852 GCTGGGCTCTGCGCATCTTCCCTGCTGTGCCCTGTGCCCTGCTCCTCCTGCTGCTCCCTTG 911
QY	915 CTGAGGACTGGGGTCTCCCTCCTGCTGATGAGGCTCTCCTACTGCTCTCTCTCCCTTG 974
Db	912 CTGAGGACTGGGGTCTCCCTCCTGCTGATGAGGCTCTCCTACTGCTCTCTCCCTTG 971
QY	975 CACAGGAAATGACAGGGGCGAGGACTTGGGAGGTTCCAGCAGCGGGGAGCCCGCCAGCA 1034
Db	972 CACAGGAAATGACAGGGGCGAGGACTTGGGAGGTTCCAGCAGCGGGGAGCCCGCCAGCA 1031
QY	1035 GTGGGACACTCTCCCTCCCGCAGTGAGCAGAGGACCCGCTGGGCTCTGCGCCCTT 1094
Db	1032 GTGGGACACTCTCCCTCCCGCAGTGAGCAGAGGACCCGCTGGGCTCTGCGCCCTT 1091
QY	1095 GCTGAGTGAGCCACCTTGTGCAACATTTGTTCTGAGGGGCCAAAGAGATGGCGTCTT 1154
Db	1092 GCTGAGTGAGCCACCTTGTGCAACATTTGTTCTGAGGGGCCAAAGAGATGGCGTCTT 1151
QY	1155 GGTCAATTTGCCCGCATGGTTGGCAGTTGGTGTAGGCCATGACAGAACTTACGGTAACA 1214
Db	1152 GGTCAATTTGCCCGCATGGTTGGCAGTTGGTGTAGGCCATGACAGAACTTACGGTAACA 1211
QY	1215 GGCACGGCTGGCCCAATGCTGCTGGAGCTGGAGCTTGCTTTGGCTTTTCAAGTGGG 1274
Db	1212 GGCACGGCTGGCCCAATGCTGCTGGAGCTGGAGCTTGCTTTGGCTTTTCAAGTGGG 1271
QY	1275 CTCGTGACGTACAGCCAGCGCGCTGCTCCTCATCTCAGTCTAGGGGGCAGCAGATATG 1334
Db	1272 TCCGTGACGTACAGCCAGCGCGCTGCTCCTCATCTCAGTCTAGGGGGCAGCAGCATAT 1331
QY	1335 GGG 1337
Db	1332 GGG 1334
RESULT 2	
AK023914	
LOCUS	2331 bp mRNA linear PRI 01-AUG-2002
DEFINITION	Homo sapiens cDNA FLJ13852 fis, clone THYR0100934, weakly similar to PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).
ACCESSION	AK023914
VERSION	AK023914.1 GI:10435995
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Tanimoto,J., Wakamatsu,A., Nakamura,Y., Nagaharai,K., Masuno,Y. and Oshima,A. NEDO human cDNA sequencing project
TITLE	Unpublished
JOURNAL	2 (bases 1 to 2331)
REFERENCE	Isogai,T. and Otsuki,T.
AUTHORS	Direct Submission
TITLE	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
JOURNAL	

Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hrl.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; CDNA full insert sequencing; Research Association for Biotechnology; CDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

## FEATURES

```

source
1 2331
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYRO100934"
/tissue_type="thyroid gl
/clone_lib="THYRO1
/note="cloning vector: p
27..851
CDS

```

```

CDS
27..851
/note="unnamed protein product"
/codon_start=1

```

```

/coord_start="BAB14721.1"
/protein_id="BAB14721.1"
/db_xref="GI:10435996"
/translation="MAAEPSRRVFGAGRMAGIAQGLIRAGKVAQHIILASAPT
DRNLCHFOALGRTTSHNQVLSCLLVIFATKPHVLPVLAEAVPVTVTHILVSVAV
AGVSLTELLLPNTFVRLVPLPCVQEGAIWMARGRHVGSSETKLLQHLLEACG
RCEVPYAVDIIHTGSGVGVAFSEALGAGVKMGPSLSAHRHTAATQGLILGTAG
MLLHGQHPQLKQSDVCTPGGTTIYGLHLEQGGRLAATMTSAVEATACRAKELSRK

```

BASE COUNT	398 a	753 c	748 g	432 t
BASE COUNT	398 a	753 c	748 g	432 t

ORIGIN	END COUNT	0000	0001	0002	0003	0004	0005	0006	0007	0008	0009	0010	0011	0012	0013	0014	0015	0016	0017	0018	0019	0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	0080	0081	0082	0083	0084	0085	0086	0087	0088	0089	0090	0091	0092	0093	0094	0095	0096	0097	0098	0099
0000	0001	0002	0003	0004	0005	0006	0007	0008	0009	0010	0011	0012	0013	0014	0015	0016	0017	0018	0019	0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	0080	0081	0082	0083	0084	0085	0086	0087	0088	0089	0090	0091	0092	0093	0094	0095	0096	0097	0098	0099		

Query Match 97.7%: Score 1307: DB 9: Length 2331:

Query Match 97.7%, Score 1307, DB 3, Length 2331,  
Best Local Similarity 99.2%: Pred. No. 1e-224:

BEST LOCAL SIMILARITY 35.2%, FREQ. NO. 1E-224,  
Matches 1313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY	15	GTCCAGGCAACAGATGGCAGCTGCGAGACCGCTCTCCGGGGCGGTGGGCTTCGTGGGC	74
DB	12	GTCCAGAGCAACAGATGGCAGCTGCGAGACCGCTCTCCGGGGCGGTGGGCTTCGTGGGC	71
QY	75	CGGGGCCGATGCGGGGGCCATCGCCAGGGCTCATCAGACGAGAAAAGTGGAACT	134
DB	72	CGGGGCCGATGCGGGGGCCATCGCCAGGGCTCATCAGACGAGAAAAGTGGAACT	131
QY	135	CAGCACATACTGGCCAGTGCAACACAGACAGAACTATGTCACITTTCAAGCTCTGGGT	194
DB	132	CAGCACATACTGGCCAGTGCAACACAGACAGAACTATGTCACITTTCAAGCTCTGGGT	191
QY	195	TGCGGACACAGCACTCCCAACAGAGGCTGTCAGAGCTGCTGCTCATCTTTGCC	254
DB	192	TGCGGACACAGCACTCCCAACAGAGGCTGTCAGAGCTGCTGCTCATCTTTGCC	251
QY	255	ACCAAGCCTCATGTGTCGCCAGCTGTCTTGGCAGAGTGCTGCTGTGTACCACACTGAA	314
DB	252	ACCAAGCCTCATGTGTCGCCAGCTGTCTTGGCAGAGTGCTGCTGTGTACCACACTGAA	311
QY	315	CACATCTTGGTGTCCGTGGCTGTGGGGTGTCTCTGAGCACCTGAGAGAGCTGCTGCC	374
DB	312	CACATCTTGGTGTCCGTGGCTGTGGGGTGTCTCTGAGCACCTGAGAGAGCTGCTGCC	371
QY	375	CCAAACACACGGGTGCTCGGGTCTTGCCAACTGTCCTGTGTGTCCAGGAAGGGCC	434
DB	372	CCAAACACACGGGTGCTCGGGTCTTGCCAACTGTCCTGTGTGTCCAGGAAGGGCC	431
QY	435	ATAGTGATGGCGGGGGCCCGCCAGTGGGAGCAGCAGACCAAGCTCTGTGAGCATCTG	494
DB	432	ATAGTGATGGCGGGGGCCCGCCAGTGGGAGCAGCAGACCAAGCTCTGTGAGCATCTG	491
QY	495	CTGGAGGCTGTGGCGGTGTGAGAGTGCTTGAAGCTAGCTGCACATCCACACTGCG	554
DB	492	CTGGAGGCTGTGGCGGTGTGAGAGTGCTTGAAGCTAGCTGCACATCCACACTGCG	551
QY	555	CTCAGTGGCAGTGGCGTGGCCCTTCGTGTGTGCATTTCTCCGAGGCCCTGGCTGAAGGACC	614

PN JP 2002191363-A/11205  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAO  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS Location/Qualifiers  
1..2615  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 492 a 831 c 818 g 474 t  
ORIGIN  
Query Match 96.5%; Score 1291.8; DB 6; Length 2615;  
Best Local Similarity 99.1%; Pred. No. 5.3e-222;  
Matches 1299; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 27 AGATGGCAGCTGCGAGCGCTCTCCGCGCGCGTGGCTTCGTGGCGCGCGCGCGCATG 86  
DB 2615 AGATGGCAGCTGCGAGCGCTCTCCGCGCGCGTGGCTTCGTGGCGCGCGCGCGCATG 2556  
QY 87 GCGGGGCCATCGCGAGCGCTCTATCAGAGCAGAGAAAGTGAAGCTCAGCACATCTG 146  
DB 2555 GCGGGGCCATCGCGAGCGCTCTATCAGAGCAGAGAAAGTGAAGCTCAGCACATCTG 2496  
QY 147 GCAGTGCACCAACACAGACAGAACCTATGTCACTTTCAAGCTCTGGTTCGCGGACACG 206  
DB 2495 GCAGTGCACCAACACAGACAGAACCTATGTCACTTTCAAGCTCTGGTTCGCGGACACG 2436  
QY 207 CACTCAACACAGAGGTGTGCAGAGCTGCCTGCTGCTATCTTTGCCACCAAGCCTCAT 266  
DB 2435 CACTCAACACAGAGGTGTGCAGAGCTGCCTGCTGCTATCTTTGCCACCAAGCCTCAT 2376  
QY 267 GTGCTGCCAGCTGCTCTGCAGAGTGGCTCTGTGGTTCACCACTGAACACATCTTGGT 326  
DB 2375 GTGCTGCCAGCTGCTCTGCAGAGTGGCTCTGTGGTTCACCACTGAACACATCTTGGT 2316  
QY 327 TCCGTGGCTGCTGGGTGTCTCTGAGCACCTCTGGAGAGCTGCTGCCCCCAACACACAGG 386  
DB 2315 TCCGTGGCTGCTGGGTGTCTCTGAGCACCTCTGGAGAGCTGCTGCCCCCAACACACAGG 2256  
QY 387 GTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 446  
DB 2255 GTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2196  
QY 447 CGGGCGCGCAGCTGGGAGCAGCAGACCAAGCTCTCTGAGCATCTCTGGAGCGCTGT 506  
DB 2195 CGGGCGCGCAGCTGGGAGCAGCAGACCAAGCTCTCTGAGCATCTCTGGAGCGCTGT 2136  
QY 507 GGGCGGTGTGAGAGGTGCCTGAAGCCTACGTTCGACATCCACACTGGCCTCAGTGGCAGT 566  
DB 2135 GGGCGGTGTGAGAGGTGCCTGAAGCCTACGTTCGACATCCACACTGGCCTCAGTGGCAGT 2076  
QY 567 GCGTGGCTTCTGT 626  
DB 2075 GCGTGGCTTCTGT 2016  
QY 627 ATGCCACAGCCTGGCCACCGCATCGTCCAGACCCCTGCTGGGAGCGGCCAAGATG 686  
DB 2015 ATGCCACAGCCTGGCCACCGCATCGTCCAGACCCCTGCTGGGAGCGGCCAAGATG 1956  
QY 687 CTGCTGCAGAGGGCCACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTGGC 746  
DB 1955 CTGCTGCAGAGGGCCACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTGGC 1896

QY 747 ACCACCATCTATGACTCCACGCCCTGGAGCAGGGCGGCTGCGAGCAGCCACCATGAGC 806  
DB 1895 ACCACCATCTATGACTCCACGCCCTGGAGCAGGGCGGCTGCGAGCAGCCACCATGAGC 1836  
QY 807 GCCGTGGAGGCTGCACCTGCCGGGCCAAGAGCTCAGCAGAAAGTAGGTAGGCTGGCTCTGG 866  
DB 1835 GCCGTGGAGGCTGCACCTGCCGGGCCAAGAGCTCAGCAGAAAGTAGGTAGGCTGGCTCTGG 1776  
QY 867 CCATCTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926  
DB 1775 CCATCTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1716  
QY 927 GCTCCCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986  
DB 1715 GCTCCCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1656  
QY 987 CAGGGGCGAGACTTTGGAGGTTCCAGCAGCGGGGAGCGCCGACAGTGGGGACACATC 1046  
DB 1655 CAGGGGCGAGACTTTGGAGGTTCCAGCAGCGGGGAGCGCCGACAGTGGGGACACATC 1596  
QY 1047 CTCCTCCCGACGTGAGCAGAGGACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1106  
DB 1595 CTCCTCCCGACGTGAGCAGAGGACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1536  
QY 1107 CCACCTTGTCTGCAACATTTGTTCTGAGGGGCCCAAGAGATGGCTGCTTGGTCTATTTGCC 1166  
DB 1535 CCACCTTGTCTGCAACATTTGTTCTGAGGGGCCCAAGAGATGGCTGCTTGGTCTATTTGCC 1476  
QY 1167 GCATGTTGGGCGAGTTGGTTGAGGCCATGAACAGAACTTACGGTAACAGACACGGCTGGC 1226  
DB 1475 GCATGTTGGGCGAGTTGGTTGAGGCCATGAACAGAACTTACGGTAACAGACACGGCTGGC 1416  
QY 1227 CCAATGCTCTGCTGGAGCTGGAGCTTGGCTTTGGCTTTCCAAAGTGGCTGCTGCGACGTA 1286  
DB 1415 CCAATGCTCTGCTGGAGCTGGAGCTTGGCTTTGGCTTTCCAAAGTGGCTGCTGCGACGTA 1356  
QY 1287 CAGCCAGCGCGGCTGCTCATCTCAGCTTAGGGGCGACGACATATGGGG 1337  
DB 1355 CAGCCAGCGCGGCTGCTCATCTCAGCTTAGGGGCGACGACATATGGGG 1305  
RESULT 4  
AK001500/c  
LOCUS Homo sapiens cDNA FLJ10638 fis, clone NT2RP2005690.  
DEFINITION AK001500  
ACCESSION AK001500  
VERSION AK001500.1 GI:7022794  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K.,  
Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y.,  
Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K.,  
Masuko, Y., and Kanehori, K.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2615)  
AUTHORS Isogai, T. and Otsuki, T.  
TITLE Direct Submission  
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5' & 3'-end one pass sequencing and clone selection:  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,



gene	1. .2400									
	/gene="DKFzp761H0716"									
CDS	<1. .735									
	/note="DKFzp761H0716"									
	/gene="Similarity to pyrroline 5-carboxylate reductase isoform (Homo sapiens)"									
	/codon_start=1									
	/product="hypothetical protein"									
	/protein_id="CAD38716.1"									
	/db_xref="GI:21739341"									
	/translation="GKVEAOHILASAPTDRNLCHFQALGCRTHSNQEVLOSCLLWIF									
	ATRPHVLPAVLAEPVPTTEHLIVSVAAGVSLTDELLPPNTRVLRLVLPNLPVVQ									
	EGALVMARGHVSSEPNLLOHLEACGCEEPYEAVIDHTGLSGSVAFAVCAFSEA									
	LAEGVAKMPSLSLAHRIAQTLLGTAKMLLHEGHPAQLRSDVCTPGGTTIYGLHAL									
	EQGLRAATMSAVEAATCRAKELSRK"									
	2381									
polyA_site	/gene="DKFzp761H0716"									
BASE COUNT	452 a	756 c	740 g	452 t						
ORIGIN										
	Query Match 89.7%; Score 1200.4; DB 9; Length 2400;									
	Best Local Similarity 99.1%; Pred. No. 1.3e-205;									
	Matches 1207; Conservative 0; Mismatches 11; Indels 0; Gaps 0;									
Qy	120	GGAAAAGTGGAAAGCTCAGCACATACTGGCCAGTGCACCAACAGACAGAAACCTATGTCCAC	179							
Db	1	GGAAAAGTGGAAAGCTCAGCACATACTGGCCAGTGCACCAACAGACAGAAACCTATGTCCAC	60							
Qy	180	TTTCAAGCTCTGGGTTGCCGGACCAACGACATCCAAACAGGAGGTGTCAGAGCTGCCTG	239							
Db	61	TTTCAAGCTCTGGGTTGCCGGACCAACGACATCCAAACAGGAGGTGTCAGAGCTGCCTG	120							
Qy	240	CTCGTCATCTTTGGCCACCAAGCCCTCATGTGCTGCAGCTGCTCTGGCCAGAGGTGGCTCCT	299							
Db	121	CTCGTCATCTTTGGCCACCAAGCCCTCATGTGCTGCCAGCTGCTCTGGCCAGGAGTGGCTCCT	180							
Qy	300	GTGGTCCACCACTGAACACATCTTTGGTGTCCGTGGCTGGGTGTCTCTGAGCACCCCTG	359							
Db	181	GTGGTCCACCACTGAACACATCTTTGGTGTCCGTGGCTGGGTGTCTCTGAGCACCCCTG	240							
Qy	360	GAGGAGCTGTGCCCCAAACACACAGGGTGCCTGGGGTCTTGTCCCAACCTGCCCTGTGTG	419							
Db	241	GAGGAGCTGTGCCCCAAACACACAGGGTGCCTGGGGTCTTGTCCCAACCTGCCCTGTGTG	300							
Qy	420	GTCACGAAAGGGCCATAGTGATGGCGCGGGCGCCACGCTGGGGAGCAGCGAGACCAAG	479							
Db	301	GTCACGAAAGGGCCATAGTGATGGCGCGGGCGCCACGCTGGGGAGCAGCGAGACCAAC	360							
Qy	480	CTCCTGCAGCATCTGCTGGAGGCCTGTGGCGGTGTGTGAGGAGTGCCTGAAGCCTACGTC	539							
Db	361	CTCCTGCAGCATCTGCTGGAGGCCTGTGGCGGTGTGTGAGGAGTGCCTGAAGCCTACGTC	420							
Qy	540	GACATCCACACTGGCCCTCAGTGGCAGTGGCGTGGCCCTCTGTGTGTGCATTCTCCGAGGCC	599							
Db	421	GACATCCACACTGGCCCTCAGTGGCAGTGGCGTGGCCCTCTGTGTGTGCATTCTCCGAGGCC	480							
Qy	600	CTGGCTGAAGGAGCCGTCAAGATGGGCATCCCGACAGCCCTGGCCACCCGATCGCTGCC	659							
Db	481	CTGGCTGAAGGAGCCGTCAAGATGGGCATCCCGACAGCCCTGGCCACCCGATCGCTGCC	540							
Qy	660	CAGACCTGTGGGGACGGCCAAAGATGCTGCTGCACGAGGGCCCAACACCCAGCCCAAGCTG	719							
Db	541	CAGACCTGTGGGGACGGCCAAAGATGCTGCTGCACGAGGGCCCAACACCCAGCCCAAGCTG	600							
Qy	720	CGCTCAGACGTGTGCACCCCGGTGGGACCAACATCTATGGATTCACGCGCCTGGAGCAG	779							
Db	601	CGCTCAGACGTGTGCACCCCGGTGGGACCAACATCTATGGATTCACGCGCCTGGAGCAG	660							
Qy	780	GGCGGGCTGGAGCAGCCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGGAG	839							
Db	661	GGCGGGCTGGAGCAGCCACCATGAGCGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGGAG	720							
Qy	840	CTCAGCAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCTCCTGCCCTGTGCCCCCTGCCCTCTC	899							



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 23 Row: 1 Column: 8  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10435995.

## FEATURES

source

Location/Qualifiers  
1..1178  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="LocusID:65263"  
/db\_xref="taxon:9606"  
/clone="MGC:14985 IMAGE:3533609"  
/tissue\_type="Lung, small cell carcinoma"  
/clone\_lib="NIH\_MGC\_7"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
47..871  
/codon\_start=1  
/product="hypothetical protein FLJ13852"  
/protein\_id="AAH07993.1"  
/db\_xref="GI:14124940"  
/translation="MAAAPPSPRRYFVGARMAGIAQGLIRAKVBAQHILASAPT  
DRNLCHFALGCRTHSNQEVLSCLLVIFATKPHVLAFLAEPVVTTHILYSVA  
AGVSLTLEELPRTVRLVLPNLPVVOEGAIVMKGRHVGSSETNLLQHLLEACG  
RCVEPEAVDILHTGLSGGVAFVCAFEALAEAGAVKMGMPSSLAHRAOTLLGTAK  
MLLHGOHPAQLRSVCPGPTTIYGLHLEOGGLRAATMSAVERATCRKELSRK"  
211 a 379 c 383 g 205 t

## CDS

BASE COUNT

ORIGIN

Query Match 84.3%; Score 1127.8; DB 9; Length 1178;  
Best Local Similarity 99.8%; Pred. No. 1.7e-192;  
Matches 1129; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 15 GTCGAGGCAACAAGATGCGAGCTGCGGAGCCGCTCTCCGCGGGCTTCGTGGGC 74  
DB 32 GTCGAGGCAACAAGATGCGAGCTGCGGAGCCGCTCTCCGCGGGCTTCGTGGGC 91  
QY 75 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCT 134  
DB 92 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCT 151  
QY 135 CAGCACATAGTGGCCAGTGCCACCAACAGACAGAGAACTATGTCACTTTCAAGCTCTGGGT 194  
DB 152 CAGCACATAGTGGCCAGTGCCACCAACAGACAGAGAACTATGTCACTTTCAAGCTCTGGGT 211  
QY 195 TGGCGGACACGACACTCCAAACAGAGAGTGCTGCAGAGCTGCCTGCTCATCTTTGCC 254  
DB 212 TGGCGGACACGACACTCCAAACAGAGAGTGCTGCAGAGCTGCCTGCTCATCTTTGCC 271  
QY 255 ACCAAGCCTCATGTGCTGCCAGCTGCTCGGACAGAGTGCTGCTGTGTCACTCACTGAA 314  
DB 272 ACCAAGCCTCATGTGCTGCCAGCTGCTCGGACAGAGTGCTGCTGTGTCACTCACTGAA 331  
QY 315 CACATCTTGCTGCTCGGCTGCTGCGGTGTCTCTGAGCACCCCTGGAGGAGCTGTGCC 374  
DB 332 CACATCTTGCTGCTCGGCTGCTGCGGTGTCTCTGAGCACCCCTGGAGGAGCTGTGCC 391  
QY 375 CCAACACACGGGTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGCTCAGGAGGGCC 434  
DB 392 CCAACACACGGGTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGCTCAGGAGGGCC 451  
QY 435 ATAGTGATGCGGGGGCCCGCAGTGGGAGCAGGACCAAGCTCTCTGCAAGCATCTG 494  
DB 452 ATAGTGATGCGGGGGCCCGCAGTGGGAGCAGGACCAAGCTCTCTGCAAGCATCTG 511  
QY 495 CTGAGGCGCTGTGGCGGTGTGAGGAGGTGCCTGAAGCCTACGTCGACATCCACACTGGC 554  
DB 512 CTGAGGCGCTGTGGCGGTGTGAGGAGGTGCCTGAAGCCTACGTCGACATCCACACTGGC 571  
QY 555 CTCAGTGGCAGTGGCGCTTCGTGTGTGTGCAATCTCCGAGGCCCTGGCTGAAGAGCC 614  
DB 572 CTCAGTGGCAGTGGCGCTTCGTGTGTGTGCAATCTCCGAGGCCCTGGCTGAAGAGCC 631

QY 615 GTCAGATGGCATGCGCCAGCAGCTGGCCCGCAGCATCGTGCACAGACCCCTGCTGGGG 674  
DB 632 GTCAGATGGCATGCGCCAGCAGCTGGCCCGCAGCATCGTGCACAGACCCCTGCTGGGG 691  
QY 675 ACGCCCAAGATGCTGCTGCACAGGCGCAACACACCCAGCTGCGCTCAGAGCTGTC 734  
DB 692 ACGCCCAAGATGCTGCTGCACAGGCGCAACACACCCAGCTGCGCTCAGAGCTGTC 751  
QY 735 ACCCGGGTGGCACCACCATCTATGAGCTCACGCCCTGGAGCAGGCGGGCTGCGAGCA 794  
DB 752 ACCCGGGTGGCACCACCATCTATGAGCTCACGCCCTGGAGCAGGCGGGCTGCGAGCA 811  
QY 795 GCCACCATGAGCGCGCTGGAGGCTGCCACCTGCGCGGCCAAGGAGCTCAGCAGAACTAG 854  
DB 812 GCCACCATGAGCGCGCTGGAGGCTGCCACCTGCGCGGCCAAGGAGCTCAGCAGAACTAG 871  
QY 855 GCTGGGCTGCTGGCCATCCTTTCTGCTCTGTGTGCCCTGCTCTCTGCTGTGCTTCCC 914  
DB 872 GCTGGGCTGCTGGCCATCCTTTCTGCTCTGTGTGCCCTGCTCTCTGCTGTGCTTCCC 931  
QY 915 CTGAGGACTGCGGCTCCTCTCCTCTCATGAGGCTCTCTCTCTCTCTCTCTCTCTCT 974  
DB 932 CTGAGGACTGCGGCTCCTCTCCTCTCATGAGGCTCTCTCTCTCTCTCTCTCTCTCT 991  
QY 975 CACAGGAAATGACAGGGGCGCAGGACTTTGGGAGGTTCCAGCAGGCGGGGAGCCCGACCA 1034  
DB 992 CACAGGAAATGACAGGGGCGCAGGACTTTGGGAGGTTCCAGCAGGCGGGGAGCCCGACCA 1051  
QY 1035 GTGGGACACT 1094  
DB 1052 GTGGGACACT 1111  
QY 1095 GCTCAGTACGCCACCTTGTCTGCAACATTTGTTCTGAGGGGCCCAAGAA 1145  
DB 1112 GCTCAGTACGCCACCTTGTCTGCAACATTTGTTCTGAGGGGCCCAAGAA 1162  
RESULT 7  
BD150492 794 bp DNA linear PAT 17-JAN-2003  
LOCUS Primer for synthesizing full-length cDNA and use thereof.  
DEFINITION  
BD150492  
ACCESSION  
BD150492.1 GI:27856250  
VERSION  
JP 2002191363-A/5335.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 794)  
AUTHORS  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE  
Primer for synthesizing full-length cDNA and use thereof  
JOURNAL  
Patent: JP 2002191363-A 5335 09-JUL-2002;  
COMMENT  
HELIOS RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002191363-A/5335  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC  
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT source 1..794  
/organism="Homo sapiens"  
/organism="Homo sapiens (human)".  
FEATURES  
source  
1..794  
/organism="Homo sapiens"



Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 27, 2003 this sequence version replaced gi:28261622.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L23062  
 Center clone name: 118\_P\_12  
 -----

Only the middle 101.8 kilobases of this clone are being submitted.  
 The remainder overlaps either accession number AC067930 [WICGR project L10100]  
 or accession number AC105219 [WICGR project L23081].

#### FEATURES

source  
 Location/Qualifiers  
 1..101777  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8"  
 /clone="RP11-118P12"  
 /clone.lib="RPC1-11 Human Male BAC"  
 1437..1459  
 /rpt\_family="GC-rich"  
 /rpt\_family="(CCCCCA)n"  
 3950..4302  
 /rpt\_family="MSTA"  
 6401..6583  
 /rpt\_family="G-rich"  
 8806..8835  
 /rpt\_family="GC-rich"  
 8878..8900  
 /rpt\_family="GC-rich"  
 13594..13939  
 /rpt\_family="MLT1B"  
 14312..14755  
 /rpt\_family="MLT1D"  
 14755..14875  
 /rpt\_family="MLT1D-int"  
 14876..15169  
 /rpt\_family="AluSg"  
 15304..15357  
 /rpt\_family="MLT1D-int"  
 15399..15698  
 /rpt\_family="MLT1D-int"  
 15716..16013  
 /rpt\_family="MLT1D"  
 16922..16957  
 /rpt\_family="AT-rich"  
 17644..17691  
 /rpt\_family="Charliel"  
 17717..17926  
 /rpt\_family="Charliel"  
 complement(17927..18232)  
 /rpt\_family="AluSx"  
 18233..18299  
 /rpt\_family="Charliel"  
 20895..20951  
 /rpt\_family="Alu"  
 complement(20965..21069)

repeat\_region  
 /rpt\_family="MER53"  
 21227..21518  
 /rpt\_family="AluY"  
 complement(23096..23135)  
 /rpt\_family="MIR3"  
 complement(25519..25676)  
 /rpt\_family="AluSc"  
 complement(25677..25973)  
 /rpt\_family="AluSx"  
 complement(25974..26107)  
 /rpt\_family="AluSc"  
 28072..28094  
 /rpt\_family="GC-rich"  
 28328..28365  
 /rpt\_family="GC-rich"  
 28373..28388  
 /note="single clone coverage"  
 28382..28386  
 /note="<30 qual single clone coverage"  
 28473..28500  
 /rpt\_family="GC-rich"  
 28492..28593  
 /note="single clone coverage"  
 28503..28507  
 /note="<30 qual single clone coverage"  
 28510..28554  
 /rpt\_family="GC-rich"  
 28569..28599  
 /rpt\_family="(CCCCCG)n"  
 28604..28640  
 /rpt\_family="GC-rich"  
 29071..29194  
 /rpt\_family="GC-rich"  
 29264..29321  
 /rpt\_family="GC-rich"  
 31816..31845  
 /rpt\_family="(TA)n"  
 32414..32714  
 /rpt\_family="AluSx"  
 complement(34498..34792)  
 /rpt\_family="AluSg"  
 38000..38153  
 /rpt\_family="(CA)n"  
 38154..38184  
 /rpt\_family="(CAGC)n"  
 38185..38211  
 /rpt\_family="(CA)n"  
 38278..38464  
 /rpt\_family="(CA)n"  
 Query Match 48.8%; Score 653.2; DB 9; Length 101777;  
 Best Local Similarity 98.18; Pred. No. 9.1e-108;  
 Matches 661; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
 QY 664 CCCTGCTGGGACGGCCCAAGATGCTGTCACGAGGGCCACACCCAGCCAGCTGCCT 723  
 Db 36491 CTCCTCAGGGGACGGCCCAAGATGCTGTCACGAGGGCCACACCCAGCCAGCTGCCT 36432  
 QY 724 CAGACGTGTGCACCCCGGGTGGCCACCACCATCTATGACTCCAGCCCTGGAGCGGGC 783  
 Db 36431 CAGACGTGTGCACCCCGGGTGGCCACCACCATCTATGACTCCAGCCCTGGAGCGGGC 36372  
 QY 784 GGCTGCGAGCAGCCACCATGAGCCCGTGGAGGCTGCCACTGCCGGGCCCAAGAGCTCA 843  
 Db 36371 GGCTGCGAGCAGCCACCATGAGCCCGTGGAGGCTGCCACTGCCGGGCCCAAGAGCTCA 36312  
 QY 844 GCAGAAAGTAGGCTGGGCTCTGGCCATCTCTTTCTGCTCTGTGCTGCTGCTGCTGCTGCTG 903  
 Db 36311 GCAGAAAGTAGGCTGGGCTCTGGCCATCTCTTTCTGCTCTGTGCTGCTGCTGCTGCTG 36252  
 QY 904 TGTCCCTTCCCTCAGGACTCGGCTCCCTCCCTCCCTGTCATGAGGGTCTCCTACTGCTCC 963  
 Db 36251 TGTCCCTTCCCTCAGGACTCGGCTCCCTCCCTCCCTGTCATGAGGGTCTCCTACTGCTCC 36192

QY 964 TTCTCCCTTGCACAGGAAATGCAGGGGCGAGGACTTGGAGGTTCCAGCAGGCGGGG 1023  
|||||  
Db 36191 TTCTCCCTTGCACAGGAAATGCAGGGGCGAGGACTTGGAGGTTCCAGCAGGCGGGG 36132  
QY 1024 AGCCCGACAGGAGGACACCTCTCCCTCCCGAGTGCAGCAGGAGCCGTGGTGGTGG 1083  
|||||  
Db 36131 AGCCCGACAGGAGGAGGACACCTCTCCCTCCCGAGTGCAGCAGGAGCCGTGGTGGTGG 36072  
QY 1084 CTCTGCCCTTGTCTGAGTGAGCCACCTTGTCTGCAACATTTGTTGAGGGGCCCAAGA 1143  
|||||  
Db 36071 CTCTGCCCTTGTCTGAGTGAGCCACCTTGTCTGCAACATTTGTTGAGGGGCCCAAGA 36012  
QY 1144 GATGGCTTGTGTCATTTGCCCGCATGTTGGGCGAGTGTGGAGCCATGAACAGAAC 1203  
|||||  
Db 36011 GATGGCTTGTGTCATTTGCCCGCATGTTGGGCGAGTGTGGAGCCATGAACAGAAC 35952  
QY 1204 TTACGGTAACAGCAGCGCTGCCCAATGCTGTCGAGTGTGAGCTGAGCTTGGCTTGGCT 1263  
|||||  
Db 35951 TTACGGTAACAGCAGCGCTGCCCAATGCTGTCGAGTGTGAGCTGAGCTTGGCTTGGCT 35892  
QY 1264 TTCCAGTGGCTCTGTCAGCTACAGCCAGGCGGCTGCCTCATCTCAGCTTAGGGGGC 1323  
|||||  
Db 35891 TTCCAGTGGCTCTGTCAGCTACAGCCAGGCGGCTGCCTCATCTCAGCTTAGGGGGC 35832  
QY 1324 ACAGCATATGGG 1337  
|||||  
Db 35831 ACAGCATATGGG 35818

RESULT 9  
AC087823/C  
LOCUS  
DEFINITION Homo sapiens chromosome 8 clone RP11-545C16 map 8, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 70 unordered pieces.  
AC087823  
VERSION AC087823.2 GI:13357313  
KEYWORDS HTG: HTGS, PHASE1  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 164959)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,  
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,  
Gardyna,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,  
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,  
Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,  
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,  
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,J., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
TITLE Submitted (28-JAN-2001) Whitehead Institute/MIT Center for Genome  
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Mar 16, 2001 this sequence version replaced gi:12584321.  
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L11736  
Center clone name: 545\_C\_16  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 70 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 489: contig of 489 bp in length  
\* 490 gap of 100 bp  
\* 590 589: contig of 506 bp in length  
\* 1196 1195: contig of 506 bp in length  
\* 1296 1295: contig of 658 bp in length  
\* 1954 2053: contig of 100 bp  
\* 2734 2733: contig of 680 bp in length  
\* 2834 2833: gap of 100 bp  
\* 3423 3422: contig of 589 bp in length  
\* 4134 4133: gap of 100 bp  
\* 4234 4233: contig of 611 bp in length  
\* 5154 5153: contig of 920 bp in length  
\* 5254 5253: gap of 100 bp  
\* 6106 6105: contig of 852 bp in length  
\* 6206 6205: gap of 100 bp  
\* 7237 7236: contig of 1031 bp in length  
\* 7337 7336: gap of 100 bp  
\* 7720 7719: contig of 383 bp in length  
\* 7820 7819: gap of 100 bp  
\* 8403 8402: contig of 583 bp in length  
\* 8503 19472: contig of 10970 bp in length  
\* 19473 19572: gap of 100 bp  
\* 19573 20298: contig of 725 bp in length  
\* 20398 21307: gap of 100 bp  
\* 21308 21407: contig of 910 bp in length  
\* 21408 22443: gap of 100 bp  
\* 22444 22344: contig of 1036 bp in length  
\* 22344 23637: contig of 1093 bp in length  
\* 23637 24910: gap of 100 bp  
\* 24911 25010: contig of 1174 bp in length  
\* 25011 25940: gap of 100 bp  
\* 25941 26040: contig of 930 bp in length  
\* 26041 27243: contig of 1203 bp in length  
\* 27244 28274: gap of 100 bp  
\* 28275 28374: contig of 931 bp in length  
\* 28375 29646: gap of 100 bp  
\* 29647 29747: contig of 1272 bp in length  
\* 29747 31485: gap of 100 bp  
\* 31486 31585: contig of 1739 bp in length  
\* 31586 32456: contig of 871 bp in length  
\* 32457 32557: gap of 100 bp  
\* 32557 33732: contig of 1176 bp in length  
\* 33733 33832: gap of 100 bp  
\* 33833 35127: contig of 1295 bp in length  
\* 35128 36370: gap of 100 bp  
\* 36371 36470: contig of 1143 bp in length  
\* 36471 37453: gap of 100 bp  
\* 37454 37553: contig of 983 bp in length  
\* 37553 gap of 100 bp

\* 37554 38694: contig of 1141 bp in length  
\* 38595 38794: gap of 100 bp  
\* 38795 39930: contig of 1136 bp in length  
\* 39931 40030: gap of 100 bp  
\* 40031 41742: contig of 1712 bp in length  
\* 41743 41842: gap of 100 bp  
\* 41843 43678: contig of 1836 bp in length  
\* 43679 43778: gap of 100 bp  
\* 43779 45606: contig of 1828 bp in length  
\* 45607 47067: gap of 100 bp  
\* 47068 47863: contig of 2157 bp in length  
\* 47864 47963: gap of 100 bp  
\* 47964 50045: contig of 2082 bp in length  
\* 50046 50145: gap of 100 bp  
\* 50146 51025: contig of 880 bp in length  
\* 51026 51125: gap of 100 bp  
\* 51126 52693: contig of 1568 bp in length  
\* 52694 52793: gap of 100 bp  
\* 52794 55531: contig of 2738 bp in length  
\* 55532 55631: gap of 100 bp  
\* 55632 57290: contig of 1659 bp in length  
\* 57291 57390: gap of 100 bp  
\* 57391 58997: contig of 1607 bp in length  
\* 58998 59097: gap of 100 bp  
\* 59098 60543: contig of 1446 bp in length  
\* 60544 60643: gap of 100 bp  
\* 60644 62329: contig of 1686 bp in length  
\* 62330 62429: gap of 100 bp  
\* 62430 64069: contig of 1640 bp in length  
\* 64070 64169: gap of 100 bp  
\* 64170 66375: contig of 2206 bp in length  
\* 66376 66475: gap of 100 bp  
\* 66476 68972: contig of 2497 bp in length  
\* 68973 69072: gap of 100 bp  
\* 69073 70628: contig of 1556 bp in length  
\* 70629 70728: gap of 100 bp  
\* 70729 73435: contig of 2707 bp in length  
\* 73436 73535: gap of 100 bp  
\* 73536 75771: contig of 2236 bp in length  
\* 75772 75871: gap of 100 bp  
\* 75872 78176: contig of 2305 bp in length  
\* 78177 78276: gap of 100 bp  
\* 78277 80144: contig of 1868 bp in length  
\* 80145 80244: gap of 100 bp  
\* 80245 82363: contig of 2119 bp in length  
\* 82364 82463: gap of 100 bp  
\* 82464 84226: contig of 1763 bp in length  
\* 84227 84326: gap of 100 bp  
\* 84327 87431: contig of 3105 bp in length  
\* 87432 87531: gap of 100 bp  
\* 87532 90276: contig of 2745 bp in length  
\* 90277 90376: gap of 100 bp  
\* 90377 93023: contig of 2647 bp in length  
\* 93024 93123: gap of 100 bp  
\* 93124 95831: contig of 2708 bp in length  
\* 95832 95931: gap of 100 bp  
\* 95932 99011: contig of 3080 bp in length  
\* 99012 99111: gap of 100 bp  
\* 99112 101722: contig of 2611 bp in length  
\* 101723 101822: gap of 100 bp  
\* 101823 105034: contig of 3212 bp in length  
\* 105035 105134: gap of 100 bp  
\* 105135 108113: contig of 2979 bp in length  
\* 108114 108213: gap of 100 bp  
\* 108214 112103: contig of 3892 bp in length  
\* 112105 112205: gap of 100 bp  
\* 112206 115406: contig of 3201 bp in length  
\* 115407 115506: gap of 100 bp  
\* 115507 119714: contig of 4208 bp in length  
\* 119715 119814: gap of 100 bp  
\* 119815 125158: contig of 5344 bp in length  
\* 125159 125258: gap of 100 bp  
\* 125259 130255: contig of 4997 bp in length

\* 130256 130355: gap of 100 bp  
\* 130356 135663: contig of 5308 bp in length  
\* 135664 135763: gap of 100 bp  
\* 135764 141345: contig of 5582 bp in length  
\* 141346 141445: gap of 100 bp  
\* 141446 146381: contig of 4936 bp in length  
\* 146382 146481: gap of 100 bp  
\* 146482 155041: contig of 8560 bp in length  
\* 155042 155141: gap of 100 bp  
\* 155142 164370: contig of 9229 bp in length  
\* 164371 164470: gap of 100 bp  
\* 164471 164959: contig of 489 bp in length.  
FEATURES  
Location/Qualifiers  
Query Match 48.8%; Score 653.2; DB 2: Length 164959;  
Best Local Similarity 98.1%; Pred. No. 8.2e-108;  
Matches 661; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 664 CCTCTGCTGGGACGGCCAAAGATGCTGTGCACGAGGGCCAAACACCCAGCCAGCTGCCT 723  
Db 141067 CTCCTCAGGGAGCGGCCAAAGATGCTGTGCACGAGGGCCAAACACCCAGCCAGCTGCCT 141008  
QY 724 CAGAGTGTGACCCCGGGTGGCACCACCATCTATGACTCCACGCCCTGAGCAGGGCG 783  
Db 141007 CAGAGTGTGACCCCGGGTGGCACCACCATCTATGACTCCACGCCCTGAGCAGGGCG 140948  
QY 784 GGCTCGAGCAGCCACCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCA 843  
Db 140947 GGCTCGAGCAGCCACCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCA 140888  
QY 844 CGAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCTCCCTCTGTGCCCTCTCTCCCTG 903  
Db 140887 CGAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCTCCCTCTGTGCCCTCTCTCCCTG 140828  
QY 904 TGCTCCCTTCCCTGAGGACTGGGGTCCCTCCCTCCTCGATGAGGGTCTCTACTGCTCC 963  
Db 140827 TGCTCCCTTCCCTGAGGACTGGGGTCCCTCCCTCCTCGATGAGGGTCTCTACTGCTCC 140768  
QY 964 TTCTCCCTTGACACAGGAAATGACAGGGGAGGACTTGGGAGTTCCAGCAGGGGGGG 1023  
Db 140767 TTCTCCCTTGACACAGGAAATGACAGGGGAGGACTTGGGAGTTCCAGCAGGGGGGG 140708  
QY 1024 AGCCCCGACAGTGGGGACACTCTCTCCCTCCCTAGTGAGCAGAGGACCCCTGTGGTGG 1083  
Db 140707 AGCCCCGACAGTGGGGACACTCTCTCCCTCCCTAGTGAGCAGAGGACCCCTGTGGTGG 140648  
QY 1084 CTCTGCCCTTCTCTGAGTGAGCCACCTTGTGCAACATTGGTTCTGAGGGGGCCCAAGA 1143  
Db 140647 CTCTGCCCTTCTCTGAGTGAGCCACCTTGTGCAACATTGGTTCTGAGGGGGCCCAAGA 140588  
QY 1144 GATGGCTCTTGGTCAATTTGCCCGCATGGTTGGGAGTTGGTTGAGGGCCATGAACAGAAC 1203  
Db 140587 GATGGCTCTTGGTCAATTTGCCCGCATGGTTGGGAGTTGGTTGAGGGCCATGAACAGAAC 140528  
QY 1204 TTACGGTAACAGGCACGGCTGGCCCAATGCTTGGTGTGAGCTGAGGCTTGCCTTTGGCT 1263  
Db 140527 TTACGGTAACAGGCACGGCTGGCCCAATGCTTGGTGTGAGCTGAGGCTTGCCTTTGGCT 140468  
QY 1264 TTCCAAAGTGGGTCTGTGACGCTACAGCCAGGGCGGCTGCCCTCATCTACGCTTAGGGGGC 1323  
Db 140467 TTCCAAAGTGGGTCTGTGACGCTACAGCCAGGGCGGCTGCCCTCATCTACGCTTAGGGGGC 140408  
QY 1324 ACGAGCATATGGG 1337  
Db 140407 ACGAGCATATGGG 140394  
RESULT 10  
AC067930/c AC067930 174906 bp DNA linear PRI 24-JUL-2002  
LOCUS Homo sapiens chromosome 8, clone RP11-661A12, complete sequence.  
DEFINITION AC067930  
ACCESSION AC067930  
VERSION AC067930.7 GI:21954034

## KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 174906)

Birren,B., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 8, clone RP11-661A12

Unpublished

2 (bases 1 to 174906)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abrahams,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 174906)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckhalter,B., Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (21-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 174906)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Macdonald,P., Major,J., Matthews,C., Liu,G., Maclean,C., Macdonald,P., Meneus,L., Mlenga,V., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

## TITLE

JOURNAL

## COMMENT

Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 24, 2002 this sequence version replaced gi:21490265.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10100

Center clone name: 661\_A12

## FEATURES

Source

Location/Qualifiers

1..174906

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-661A12"

/clone\_lib="RPC1-11 Human Male BAC"

/complement(351..430)

/rpt\_family="MIR"

977..1108

/rpt\_family="L1MC4a"

/complement(1149..1449)

/rpt\_family="AluSq"

3675..3781

/rpt\_family="MIR"

/complement(16671..16974)

/rpt\_family="AluSq"

20174..20412

/rpt\_family="L2"

20487..20585

/rpt\_family="L2"

/complement(20694..20702)

/note="<30 qual SNGL region"

23655..23772

/rpt\_family="L1ME2"

24264..24383

/rpt\_family="L1ME2"

/complement(24446..24779)

/rpt\_family="MLT1b"

26300

/note="probably A; possibly G"

26681..26825

/rpt\_family="L1ME2"

28798..29091

/rpt\_family="AluX"

32096..32118

/rpt\_family="AT\_rich"

32990..33157

/rpt\_family="L1M2"

33598..33749

/rpt\_family="L1ME2"

/complement(33750..34064)

/rpt\_family="AluJo"

34065..34129

/rpt\_family="L1ME2"

34132..34706

/rpt\_family="L1MC4a"

34991..35084

/rpt\_family="L1MD3"

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat_region	35191..35325	/rpt_family="L1MC4"	143815	TGTCCTTCCCTGAGGACTGCGGCTCCCTCCCTCTGTCATGAGGCTCTCTACTGCTCC	143756
repeat_region	36901..37069	/rpt_family="CGGGG)n"	964	TTCTCCCTTCCACAGGGAATGACAGGGGCGAGGACTTGGAGGTTCACAGCGCGGGG	1023
repeat_region	37077..37162	/rpt_family="CGGGG)n"	143755	TTCTCCCTTCCACAGGGAATGACAGGGGCGAGGACTTGGAGGTTCACAGCGCGGGG	143696
repeat_region	37077..37162	/rpt_family="CGGGG)n"	1024	AGCCCGGACCACTGCGGACACTCTCTCCCTCCAGTGAGCAGAGGACCGTGGTGGTGG	1083
unsure	complement(38672..38844)	/rpt_family="MSTD"	143695	AGCCCGGACCACTGCGGACACTCTCTCCCTCCAGTGAGCAGAGGACCGTGGTGGTGG	143636
repeat_region	complement(38859..38884)	/note="<30 qual SNGL region"	1084	CTCTGCCCTTGGCTGCACTGAGCCCACTTCTGCTGCAACATTTGTTGAGGGGCCCCAAGA	1143
repeat_region	complement(38880..38983)	/rpt_family="MSTC"	143635	CTCTGCCCTTGGCTGCACTGAGCCCACTTCTGCTGCAACATTTGTTGAGGGGCCCCAAGA	143576
unsure	complement(38896..38903)	/note="<30 qual SNGL region"	1144	GATGCGCTTGGTTCATTTGCCCGCATGTTGGCGAGTTGTTGAGGCCCATGAACAAC	1203
unsure	complement(38898..38902)	/note="probably ACTGA"	143575	GATGCGCTTGGTTCATTTGCCCGCATGTTGGCGAGTTGTTGAGGCCCATGAACAAC	143516
repeat_region	complement(38984..39281)	/rpt_family="AluSx"	1204	TTACGGTAACAGGCAAGGCTGCGCCCAATGCTTGGTCTGAGCTGGAGCTGGCTTTGGCT	1263
repeat_region	complement(39282..39334)	/rpt_family="G-rich"	143515	TTACGGTAACAGGCAAGGCTGCGCCCAATGCTTGGTCTGAGCTGGAGCTGGCTTTGGCT	143456
repeat_region	complement(39366..39668)	/note="<30 qual SNGL region"	1264	TTCCAAAGTGGGCTCGTGACGCTACAGCCAGCGCGCTGCCTCATCTCAGCTCTAGGGGGC	1323
repeat_region	complement(40044..40423)	/rpt_family="AluJb"	143455	TTCCAAAGTGGGCTCGTGACGCTACAGCCAGCGCGCTGCCTCATCTCAGCTCTAGGGGGC	143396
repeat_region	complement(40423)	/rpt_family="MLTIA2"	1324	ACGACCATATGGG	1337
repeat_region	complement(41752..41933)	/rpt_family="(CCA)n"	143395	ACGACCATATGGG	143382
repeat_region	complement(42478..42510)	/rpt_family="G-rich"	143395	ACGACCATATGGG	143382
unsure	complement(43664..43676)	/note="<30 qual SNGL region"	1264	TTCCAAAGTGGGCTCGTGACGCTACAGCCAGCGCGCTGCCTCATCTCAGCTCTAGGGGGC	1323
unsure	complement(43666..43674)	/note="probably ACTGGTGGC"	143455	TTCCAAAGTGGGCTCGTGACGCTACAGCCAGCGCGCTGCCTCATCTCAGCTCTAGGGGGC	143396
unsure	complement(46481..46483)	/note="probably CCC"	1324	ACGACCATATGGG	1337
repeat_region	complement(54691..54749)	/rpt_family="(CA)n"	143395	ACGACCATATGGG	143382
repeat_region	complement(56353..56388)	/rpt_family="MER5B"	143395	ACGACCATATGGG	143382
repeat_region	complement(60486..60780)	/rpt_family="AluSx"	143395	ACGACCATATGGG	143382
repeat_region	complement(60802..61091)	/rpt_family="AluJb"	143395	ACGACCATATGGG	143382
repeat_region	complement(69122..69423)	/rpt_family="AluJb"	143395	ACGACCATATGGG	143382
repeat_region	complement(71273..71458)	/rpt_family="L2"	143395	ACGACCATATGGG	143382
repeat_region	complement(71493..71733)	/rpt_family="L2"	143395	ACGACCATATGGG	143382
repeat_region	complement(71849..72140)	/rpt_family="AluSx"	143395	ACGACCATATGGG	143382
Query Match	48.88;	Score 653.2;	DB 9;	Length 174906;	
Best Local Similarity	98.18;	Pred. No. 8.1e-108;	Mismatches 13;	Indels 0;	Gaps 0;
Matches	661;	Conservative 0;			
QY	664	CCCTGCTGGGAGGCGCAAGATGCTGTCACAGGCGCCCAACACCCAGCGCTGCCT	723		
Db	144055	CTCCTAGGGGAGCGCCCAAGATGCTGTCACAGGCGCCCAACACCCAGCGCTGCCT	143996		
QY	724	CAGAGGTGTGACCCCGGTGGGACCAACCATCTATGACTCCACGCCCTGGAGCGGGG	783		
Db	143995	CAGAGGTGTGACCCCGGTGGGACCAACCATCTATGACTCCACGCCCTGGAGCGGGG	143936		
QY	784	GGCTGCGAGGACCGCCACCATGAGCGCGTGGAGGCTGCCACTTCCCGGGCCAAAGAGCTCA	843		
Db	143935	GGCTGCGAGGACCGCCACCATGAGCGCGTGGAGGCTGCCACTTCCCGGGCCAAAGAGCTCA	143876		
QY	844	GCAGAAAGTAGGCTGGGCTCTGGCCATPCCTTTCTGCTGCTGTCGCTGCTGCTGCTG	903		
Db	143875	GCAGAAAGTAGGCTGGGCTCTGGCCATPCCTTTCTGCTGCTGTCGCTGCTGCTGCTG	143816		
QY	904	TGTCCTTCCCTGAGGACTGCGGCTCCCTCCCTCTGTCATGAGGCTCTCTACTGCTCC	963		

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC019122 207636 bp DNA linear HTG 07-JUL-2000  
Homo sapiens chromosome 8 clone RP11-545C16, WORKING DRAFT  
SEQUENCE, 12 unordered pieces.

AC019122  
AC019122.3 GI:8318560  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
Homo sapiens (human)

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Waterston, R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 207636)

Waterston, R.H.  
Direct Submission  
Submitted (30-DEC-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

On Jun 7, 2000 this sequence version replaced gi:7109639.

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Project Information

Center project name: H\_NH0545C16

Summary Statistics

Sequencing vector: M13; 88%

Chemistry: Dye-primer ET; 88% of reads

Chemistry: Dye-terminator Big Dye; 12% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 198432 bases at least Q40

Consensus quality: 201424 bases at least Q30

Consensus quality: 203070 bases at least Q20

Insert size: 182000; agarose-fp

Insert size: 206536; sum-of-contigs

Quality coverage: 6.36 in Q20 bases; agarose-fp



Quality coverage: 5.66 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 1653: contig of 1653 bp in length  
 \* 1654 1753: gap of unknown length  
 \* 1754 3687: contig of 1934 bp in length  
 \* 3688 3787: gap of unknown length  
 \* 3788 11012: contig of 7225 bp in length  
 \* 11013 11112: gap of unknown length  
 \* 11113 16540: contig of 5428 bp in length  
 \* 16541 16640: gap of unknown length  
 \* 16641 23739: contig of 7099 bp in length  
 \* 23740 23839: gap of unknown length  
 \* 23840 35039: contig of 11200 bp in length  
 \* 35040 35139: gap of unknown length  
 \* 35140 54231: contig of 19092 bp in length  
 \* 54232 54331: gap of unknown length  
 \* 54332 78381: contig of 24050 bp in length  
 \* 78382 78481: gap of unknown length  
 \* 78482 106509: contig of 28028 bp in length  
 \* 106510 106609: gap of unknown length  
 \* 106610 130451: contig of 23842 bp in length  
 \* 130452 130551: gap of unknown length  
 \* 130552 157163: contig of 26612 bp in length  
 \* 157164 157264: gap of unknown length  
 \* 157264 207636: contig of 50373 bp in length.

## FEATURES

source  
 1. .207636  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /clone="RP11-545C16"  
 misc\_feature  
 /note="assembly\_name:Contig14"  
 1754..3687  
 /note="assembly\_name:Contig15"  
 3788..11012  
 /note="assembly\_name:Contig16"  
 11113..16540  
 /note="assembly\_name:Contig17"  
 16641..23739  
 /note="assembly\_name:Contig18"  
 23840..35039  
 /note="assembly\_name:Contig19"  
 35140..54231  
 /note="assembly\_name:Contig20"  
 54332..78381  
 /note="assembly\_name:Contig21"  
 78482..106509  
 /note="assembly\_name:Contig22"  
 106610..130451  
 /note="assembly\_name:Contig23"  
 130552..157163  
 /note="assembly\_name:Contig24"  
 157264..207636  
 /note="assembly\_name:Contig25"  
 clone\_end:77  
 vector\_side:left"

BASE COUNT 44315 a 59454 c 59618 g 43137 t 11112 others  
 ORIGIN

Query Match 44.88; Score 599.4; DB 2; Length 207636;  
 Best Local Similarity 96.9%; Pred. No. 3.5e-98;  
 Matches 653; Conservative 0; Mismatches 17; Indels 4; Gaps 4;

Qy	664	CCCTGCTGGGGACGGCCAAAGATGCTGTCACAGAGGGCCAAACACCCAGCCAGCGTGGCT	723
Db	48388	CTCCTCAGGGAGGGCCAAAGATGCTGTCACAGAGGGCCAAACACCCAGCCAGCGTGGCT	48447
Qy	724	CAGACGTGTGCACCCCGGGTGGCACCACCATCTATGAGCTCCAGCCCTGAGCAGGGCG	793
Db	48448	CAGACGTGTGCACCCCGGGTGGCACCACCATCTATGAGCTCCAGCCCTGAGCAGGGCGT	48507
Qy	784	GGCTGGAGCAGCACCACATGAGCGCGTGGAGGTGCCACCTGCCGGGCCAAAGAGCTCA	843
Db	48508	GNCTGCGAGCAG-CACCATGAGCGCGTGGAGG-TGCCACCTGCCGGGCCAAAGAGCTCA	48565
Qy	844	GCAGAAAGTAGGCTGGGCTCTGGCCATCCCTTCCTGCTCTGTCCTGCTCTGCTCTGCTG	903
Db	48566	GCAGAAAGTAGGCTGGGCTCTGGG-CATCCTTTCTGCTCTGTCCTGCTCTGCTCTGCTG	48624
Qy	904	TGTCCTCTTCCCTTGAGGACTGCGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC	963
Db	48625	TGTCCTCTTCCCTTGAGGACTGCGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC	48684
Qy	964	TTCTCCCTTGCACAGGGAAATGACAGGGGGCAGGACTTGGAGGTTCACAGCAGCGGGGG	1023
Db	48685	TTCTCCCTTGCACAGGGAAATGACAGGGGGCAGGACTTGGAGGTTCACAGCAGCGGGGG	48744
Qy	1024	AGCCCGACAGTGGGGACACTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC	1083
Db	48745	AG-CCCGACAGTGGGGACACTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC	48803
Qy	1084	CTCTGCCCTTGTGTCAGTGGAGCCACCTTCTGTCGACATTTGTTCTGAGGGGCCCAAGA	1143
Db	48804	CTCTGCCCTTGTGTCAGTGGAGCCACCTTCTGTCGACATTTGTTCTGAGGGGCCCAAGA	48863
Qy	1144	GATGGGCTCTTGTGTCATTTGGCCGATGTTGGGCGAGTTGGTTGAGGCCATGAACAGAAC	1203
Db	48864	GATGGGCTCTTGTGTCATTTGGCCGATGTTGGGCGAGTTGGTTGAGGCCATGAACAGAAC	48923
Qy	1204	TTACGGTACAGCAGCGGCTGCCCAATGCTGCTGAGCTGGAGCTGGCTTGGCT	1263
Db	48924	TTACGGTACAGCAGCGGCTGCCCAATGCTGCTGAGCTGGAGCTGGCTTGGCTTGGCT	48983
Qy	1264	TTCAAAGTGGGCTCGTGCAGCTACAGCAGCGCGGCTGCCCTCATCTCAGCTCAGGGGGC	1323
Db	48984	TTCAAAGTGGGCTCGTGCAGCTACAGCAGCGCGGCTGCCCTCATCTCAGCTCAGGGGGC	49043
Qy	1324	ACGAGCATATGGG 1337	
Db	49044	ACGAGCATATGGG 49057	
RESULT 12			
LOCUS	BC026536	1276 bp mRNA linear ROD 20-SEP-2002	
DEFINITION	Mus musculus, RIKEN cDNA 1110058B13 gene, clone MGC:35759		
IMAGE	4987954, mRNA, complete cds.		
ACCESSION	BC026536		
VERSION	BC026536.1	GI:20071633	
KEYWORDS	MGC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 1276)		
JOURNAL	Strausberg, R.		
	Direct Submission		
	Submitted (02-APR-2002) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>		
COMMENT	Contact: MGC help desk		
	Email: <a href="mailto:cgabbs@mail.nih.gov">cgabbs@mail.nih.gov</a>		
	Tissue Procurement: Jeffrey E. Green, M.D.		
	CDNA Library Preparation: Life Technologies, Inc.		



CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) mcdpaxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Place: 58 Row: d Column: 3  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 13384799.

FEATURES  
source  
1..1276  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:35759 IMAGE:4987954"  
/tissue\_type="Colon, normal. 5 month old male mouse."  
/clone\_lib="NCI CGAP Co24"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
16..840  
/codon\_start=1  
/product="RIKEN cDNA 1110058B13 gene"  
/protein\_id="AAH26536.1"  
/db\_xref="GI:20071634"  
/db\_xref="locusID:66194"  
/translation="MAATMSEPRRVGVFVAGRMATIAIRGLIQKVKVAKOVLASAPT  
DNMLCFRALGQTHSNHNEVLQPLVFATKQVLTVAEVAPIVTHIIVSVA  
DNLSTMGELLPNTRVLRVPCVVGAMVARGHAGNDRELIONLEACG  
OCTLEPSEVDIHTGLSGSVAFVCTFEALAEAGAKMGPGLAHRLAQTLLGTAK  
MLQEQEKHPAQLRTDVLTPAGTTIHLHALERGGFRATMSAVEAATCRAKELSKK"  
288 a 357 c 360 g 271 t

CDS

BASE COUNT  
ORIGIN

Query Match 42.4%; Score 567.6; DB 10; Length 1276;  
Best Local Similarity 79.8%; Pred. No. 5.3e-92;  
Matches 669; Conservative 0; Mismatches 169; Indels 0; Gaps 0;  
QY 24 AACAGATGCGAGCTGGGAGCGCTCTCCGCGCGCGTGGCTTGTGGCGCGCGCGC 83  
DB 10 AGCAGGATGCGCGGCGACAATGTCTGAACCTCGGCGTGTGGCTTGTGGCTGCGAGCGCGC 69  
QY 84 ATGCGGGGGCCATCGCGAGCGCTCATCAGACGAGGAAAGTGAAGCTCAGCACATA 143  
DB 70 ATGCGGAGGCGCAFTGCCGAGGCGCTATCCAGAGGCAAGAAAGTGAAGCTTAAACAAGT 129  
QY 144 CTGGCCAGTGCCACACAGACAGAACCTATGTCACTTTCAAGCTCTGGGTTGCCGAGC 203  
DB 130 CTGGCCAGTGCCACACGACACACACCTCTGCCACTTCAGGCTCTGGGTTGCCAGACT 189  
QY 204 AGCAGCTCCACACAGGAGTGCTGCGAGAGTGGCTGTCTATCTTTGCCACCAAGCT 263  
DB 190 ACTCACTCCACACCATGAGTGCTGCAAAACTGCGCACTTGTCTATCTTTGCCACCAACC 249  
QY 264 CATGTGTGCGAGCTGCTCGGAGAGTGCTGTCTGTCTCACTTGAACACATCTTG 323  
DB 250 CAAGTCTTGCACATGCTCTGGGGAAGTGGCCCCCATATGACCTTACCTAGACATCATC 309  
QY 324 GTGTGCGGTGCTGGGCTGTCTGTGAGCAGCTGGAGAGTGTGTGCGCCCAACACA 383  
DB 310 GTATCTGTGGCTGTGGATCTCTCTGAGCACAATGGAGGGCTGTACCCCGGAACACA 369  
QY 384 CGGTGTGCTGGGCTTTGCCAACCTGCCCTGTGTGTGTCAGGAAGGGCCATAGTGATG 443  
DB 370 CGAGTATGCGAGTCTCTCCCAATCTACCTGCTGTGTCAGGAGGGGGCGCATGGTATG 429  
QY 444 GCGGGGGCGCGCAGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 503  
DB 430 GCGCGGGGCGCCACTGTGTGGGAAGGATGACGAGAGCTCTACAGAACTTGTCTGGGAAGC 489

QY 504 TGTGGCGGTGTGAGGAGTGCCTGAAGCTACGTACATCCACATCGCTGCCTCAGTGGC 563  
DB 490 TGTGGCGGTGTGAGGAGTGCCTGAAGCTACGTACATCCACATCGCTGCCTCAGTGGC 549  
QY 564 AGTGTGCTGGCTGT 623  
DB 550 AGTGTGCTGGCTGT 609  
QY 624 GGCATCCCGAGCAGCGCTGGCCACCGCATCCCGAGCCCTGCTGGGAGCGGCCAAG 683  
DB 610 GGCATCCCGAGCAGCGCTGGCCACCGCATCCCGAGCCCTGCTGGGAGCGGCCAAG 669  
QY 684 ATGTGTGCTGACGAGCGCCAAACACCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 743  
DB 670 ATGTGTGACGAGGAAGGAAGCACCACCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
QY 744 GGCATCCCGAGCAGCGCTGGCCACCGCATCCCGAGCCCTGCTGGGAGCGGCCAAG 803  
DB 730 GGCATCCCGAGCAGCGCTGGCCACCGCATCCCGAGCCCTGCTGGGAGCGGCCAAG 789  
QY 804 AGCGCGGTGGAGCTGCCACCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 861  
DB 790 AGTGTGCTGGAAGCAGCTACCTGCGCGGCTAAGGAGCTCAGCAAGAAAGTGAAGGCGC 847

RESULT 13  
BD148256 563 bp DNA linear PAT 17-JAN-2003  
LOCUS  
Primer for synthesizing full-length cDNA and use thereof.  
DEFINITION  
BD148256  
ACCESSION  
BD148256.1 GI:27854014  
VERSION  
JP 2002191363-A/3099.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 563)  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 3099 09-JUL-2002;  
JOURNAL  
HELIOS RESEARCH INSTITUTE  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002191363-A/3099  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2002080990  
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORI  
PI SAITO,  
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
PI KEIICHI NAGAI,TETSUJI OTSUKI  
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC  
10,  
PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key

FT source 1..563  
FT Location/Qualifiers  
1..563  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 99 a 171 c 181 g 108 t 4 others  
ORIGIN  
Query Match 36.4%; Score 486.6; DB 6; Length 563;  
Best Local Similarity 95.9%; Pred. No. 2.1e-77;  
Matches 540; Conservative 0; Mismatches 18; Indels 5; Gaps 4;  
QY 27 AAGATGCGAGCTGCGGAGCGCTTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 85  
|||||



DB	64055	TGGTGTGGCCCTCA	-----CTGAAGGTGCCGTCA--ATGGG	64022
QY	626	CATGCCCGACAGCCTTGGCCACCGCATCGCTGCCAGACCTGCTGGGGACGGCCAGAT		685

33

CDs

```
BASE COUNT      364 a      582 c      541 g      361 t
ORIGIN
Query Match      16.9%; Score 226.4; DB 9; Length 1848;
Best Local Similarity 58.1%; Pred. No. 8e-31;
Matches 467; Conservative 0; Mismatches 316; Indels 21; Gaps 3;
Qy 56 GCGCGTGGGCTTCGTGGGCGCGCGCATGCGGGGGCCATCGCGCAGGGGCTCATCAG 115
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 52 GAGCGTGGGCTTCATGGCGCTGCGCAGCTGGCTTTTGGCCTGGCCAGGGCTTCACAGC 111
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 116 AGCAGGAAAGTGGAGCTCAGCACATAGTGGCCAGTGACCAACACAGAGAACTATG 175
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 112 AGCAGGCGTCTTGCTGCCACAGATAATGGCTAGCTCCC---CAGACATGGACCTGGC 168
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 176 TCACCTTCAAGCTC-----TGGGTTGCCGACGACGCACTCCCAACAGGAGGTGCT 226
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 169 CACAGTTTCTGCTCTCAGGAGATGGGGTGAAGTTGACACCCCAACAAGAGACGCT 228
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 227 GCAGAGCTGCTGCTCGTCACTTTGCCAACAGCCTCATGTGCTGCCAGCTGTCCCTGGC 286
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 229 GCAGCACAGTGATGCTCTTCTGCTGCTGTGAAGCCACACATCATCCCTTCATCCTGGA 288
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 287 AGAGTGGCTTCTGTTGGTGCACACTGAACACATCTTGGTGTCCGTGGCTGTGGGGTGTG 346
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 289 TGAATAGCGCGCGACATTGAGGACAGACACATTTGTTGTTCTCGCGCGCGCGGTAC 348
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 347 TCTGAGCACCTGAGGAGCTGCTG-----CCCCCAACACACAGGGTGTGGGGT 397
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 349 CATCAGCTCCATTGAGAGAGAGCTGTACGGTTTCGGCCAGCCCCCAGGGTATCCGCTG 408
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 398 CTTGCCCAACTCCCTGTGTGTCCAGGAAGGGCCATAGTGTATGGCGCGGGCGGCCA 457
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 409 CATGACCAACACTCCAGTCTGTGCTGGGAGGGGGCCACCGTGTATGCCACAGGCACGCA 468
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 458 CTTGGGGAGCAGGAGACCAAGCTCCTGAGCATCTGTGGAGGCTGTGGGGGTGTGA 517
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 469 CGCCAGGTGGAGGAGGGAGGCTCATGGAGCAGCTGCTGAGCAGCTGGGGCTTCTGCA 528
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 518 GGAGGTGCTGAAGCCCTACGTCACATCCACACTGGCCTCAGTGGCAGTGGGCTGGCCTT 577
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 529 GGAGGTGGAAGAGGACCTGATTGATGCCGTCACGGGCTCAGTGGCAGGGCCCCCCTA 588
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 578 CGTGTGTGCATTTCCGAGGCCCTGGCTGAAGAGCCGTCAGATGGGATGCCACAGAG 637
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 589 CGCATTCACAGCCCTGGATGCCCTGTGATGGGGCGTGAAGATGGGACTTCCAAGGG 648
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 638 CCTGGCCACCCGATCGCTGCCAGACCCTGTGGGGACGGCCCAAGATGCTGTGCACGA 697
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 649 CTTGGCAGTCCGGCTCGGGGCCAGGCCCTTCTGGGGGTGCCAAGATGCTGTGCATC 708
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 698 GGGCCAAACCCAGCCCGCTGGCTCAGACGTGTGCACCCCGGGTGGCACCACTATTA 757
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 709 AGRACAGCACCCAGGCCAGCTCAAGACACAGCTCAGCTCTCTCTGGTGGGGCCACCATCA 768
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 758 TGGACTCCAGCCCTGGAGAGGGCGGGTGGCAGAGCCACCATGAGCGCGTGGAGGC 817
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 769 TGCCTTCATGTGTGGAGAGTGGGGCTTCCGCTCCCTGCTCATCAACGCTGTGGAGGC 828
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 818 TGCCACCTGCGGGCCAGGAGCT 841
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 829 CTCTGTCATCCGACACGGGAGCT 852
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 8, 2003, 16:36:16  
Job time : 4875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 338 Seconds  
(without alignments)  
8166.553 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcagctgtctgcga.....ggggcagcagcatatggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA.:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1307	97.7	2331	14	US-10-161-418A-2
2	1127.8	84.3	1178	14	US-10-161-418A-7
3	857	64.1	860	14	US-10-161-418A-6
4	226.4	16.9	1848	14	US-10-161-418A-3
5	224.8	16.8	1792	10	US-09-880-107-2400
6	224.8	16.8	1792	14	US-10-161-418A-1
7	214.4	16.0	1715	14	US-10-161-418A-5
8	211.2	15.8	1028	14	US-10-161-418A-4
9	203.4	15.2	1478	10	US-09-925-300-278
10	201.8	15.1	1708	14	US-10-161-418A-8
11	201.8	15.1	1742	9	US-09-912-717-2
12	167	12.5	999	14	US-10-161-418A-9
13	156.4	11.7	807	14	US-10-156-761-4708
14	156.4	11.7	9025608	14	US-10-156-761-1
15	129	9.6	454	11	US-09-918-995-11309
16	118.6	8.9	371	11	US-09-918-995-181

17	115.6	8.6	482	11	US-09-918-995-31035	Sequence 31035, A
18	84.4	6.3	810	10	US-09-738-626-460	Sequence 460, App
19	84.4	6.3	3309400	10	US-09-738-626-1	Sequence 1, Appli
20	83.2	6.2	415	11	US-09-918-995-17742	Sequence 17742, A
21	73.4	5.5	888	10	US-09-726-397A-4	Sequence 4, Appli
22	73.4	5.5	888	10	US-09-712-363-28	Sequence 28, Appl
23	73.4	5.5	86114	15	US-10-080-170-648	Sequence 648, App
24	72.8	5.4	931	13	US-10-027-632-161469	Sequence 161469, A
25	72.8	5.4	931	13	US-10-027-632-161470	Sequence 161470, A
26	70.6	5.3	780	10	US-09-974-300-343	Sequence 343, App
27	69.6	5.2	843	13	US-10-027-632-170626	Sequence 170626, A
28	62	4.6	855	14	US-10-128-714-2252	Sequence 2252, Ap
29	62	4.6	855	14	US-10-128-714-7252	Sequence 7252, Ap
30	59.8	4.5	1014	10	US-09-070-927A-157	Sequence 157, App
31	54.6	4.1	2908	14	US-10-128-714-252	Sequence 252, App
32	54.6	4.1	2908	14	US-10-128-714-1252	Sequence 1252, Ap
33	54.4	4.1	908	14	US-10-128-714-1252	Sequence 6252, Ap
34	54.4	4.1	908	14	US-09-726-397A-2	Sequence 2, Appli
35	53.6	4.0	2209	10	US-09-796-692-6261	Sequence 6261, Ap
36	52.2	3.9	334	10	US-10-040-862-6261	Sequence 390, App
37	52.2	3.9	334	14	US-10-156-761-390	Sequence 1531, Ap
38	48.2	3.6	2817	14	US-10-156-761-1531	Sequence 1, Appli
39	47.4	3.5	9025608	14	US-10-156-761-1	Sequence 1729, Ap
40	47.4	3.5	795	14	US-10-156-761-1729	Sequence 6088, Ap
41	47.2	3.5	1359	14	US-10-156-761-6088	Sequence 1113, Ap
42	46.4	3.5	1359	14	US-10-156-761-1113	Sequence 2474, Ap
43	45.6	3.4	834	14	US-10-156-761-2474	Sequence 807, App
44	44.4	3.3	903	14	US-10-037-270-807	
45	44	3.3	1645	14		

## ALIGNMENTS

### RESULT 1

US-10-161-418A-2  
; Sequence 2, Application US/10161418A  
; Publication No. US20030036078A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-055  
; CURRENT APPLICATION NUMBER: US/10/161.418A  
; CURRENT FILING DATE: 2002-09-11  
; PRIOR APPLICATION NUMBER: US 60/296,080  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,509  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-161-418A-2

Query Match	97.7%	Score 1307	DB 14	Length 2331
Best Local Similarity	99.2%	Pred. No. 0		
Matches 1313	Conservative	0	Mismatches 10	Indels 0
Gaps	0			
QY	15	GTCCGAGGCAACAGATGCGAGCGGCTCGCGCGCGGCTTCTGGGC	74	
Db	12	GTCCGAGGCAACAGATGCGAGCGGCTCGCGCGCGGCTTCTGGGC	71	
QY	75	CGGGCCCGCATGGCGGGGCCATCGCGAGGCGCTCATCAGAGGAGGAAAGTGAAGCT	134	
Db	72	CGGGCCCGCATGGCGGGGCCATCGCGAGGCGCTCATCAGAGGAGGAAAGTGAAGCT	131	
QY	135	CAGCACATCTGGCCAGTGCACACAGAGGACCTATGTCATCTTTCAAGCTCTGGGT	194	
Db	132	CAGCACATCTGGCCAGTGCACACAGAGGACCTATGTCATCTTTCAAGCTCTGGGT	191	
QY	195	TGCCGGACACGACATCCCAACAGAGGTGTGCAGAGCTGCCTCTCTCTCTTTC	254	







[illegible]

## RESULT 5

```

US-09-880-107-2400
; Sequence 2400, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107

```

```

; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2400
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M77836
US-09-880-107-2400

```

[illegible]



```
Db 791 CTCCTGCATCCGCACACAGGGAGCT 814
RESULT 6
US-10-161-418A-1
; Sequence 1, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-161-418A-1
Query Match 16.8%; Score 224.8; DB 14; Length 1792;
Best Local Similarity 58.0%; Pred. No. 4.2e-51;
Matches 466; Conservative 0; Mismatches 317; Indels 21; Gaps 3;
Qy 56 GCGCGTGGGCTTCGTGGGCGCGGCGCATGCGGGGGCCATCGCGAGGGCTCATCAG 115
Db 14 GAGCGTGGGCTTCATCGGCGCTGGCCAGCTGGCTTTTGGCCCTGGCCAGGGCTTCACAG 73
Qy 116 AGCAGGAAAGTGGAGCTCAGCACATCTAGCCAGTGCACCAACAGACAGAGAACCTATG 175
Db 74 AGCAGGCGTCTTGGCTGCCCAAGATAATGGCTAGCTCCC---CAGACATGGACTGGC 130
Qy 176 TCACCTTTCAAGCTC-----TGGGTTGCCGACCGACGCACTCCAAACAGGAGGTGCT 226
Db 131 CACAGTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACACCCCAACAGAGACGCT 190
Qy 227 GCAGAGTGGCTGCTCGTCATCTTTGCCAACAGCCTCATGTGCTGCCAGCTGTCTGGC 286
Db 191 GCAGCAGTGTATGTCTTCTTCTGCTGTGAAGCCACACATCATCCCCCTTCATCCGTA 250
Qy 287 AGAGTGGCTCCTGTGCTCACCCTGAACACATCTTGGTGTGCTGCTGGCTGCTGGGCTGC 346
Db 251 TGAATAGGCGCGACATTGAGGACAGACATTTGGTGTCTTGGCGCGCGCGGCTCAC 310
Qy 347 TCTGAGCACCCCTGGAGAGCTGCTG-----CCCCAAACACACGGGTGCTGGGGT 397
Db 311 CATCAGCTCCATTGAGAAGAGCTGTACAGCTTTTCCGCGACGCCGCCAGGGTCAATCGCTG 370
Qy 398 CTGGCCCAACTGCCCTGTGTGTCAGGAGGGGCCATAGTATGGCGGGGGCGGCCCA 457
Db 371 CATGACCAACACTCCAGTCTGTGTGGGGAGGGGCCACCGGTATGCCACAGACAGCA 430
Qy 458 CGTGGGAGCAGCAGACCAAGCTCCTGCAGCATCTGCTGGAGGCTGTGGGGCGGTGA 517
Db 431 CGGCCAGGTGGAGAGCGGAGGCTCATGGAGCAGCTGTGAGCAGGCTGGGCTTCTGCAC 490
Qy 518 GGAGTGCCTTGAAGCCTTACCTCGACATCCACATGGCCTCAGTGGCAGTGGCGTGGCCTT 577
Db 491 GGAGTGGAGAGGACCTGATTGATCCGCTCACGGGCTCAGTGGCAGCGGCCCGCCCTA 550
Qy 578 CGTGTGTGCAATTTCCGAGGCCCTGTGCTGAAGAGCGGCTCAAGATGGGATGCCAGCAG 637
Db 551 CGCATTCACAGCCCTGGATCCCTGTGCTGATGGGGTGTGAAGATGGGACTTCCAAGGCG 610
Qy 638 CTTGGGCCACCGCATCGCTGCCAGACCCCTGCTGGGAGCGGCCCAAGATGCTGTGCACGA 697
Db 611 CTTGGGATCTCGGCTCGGGGCCCGAGGCCCTCTTGGGGGCTGCCAAGATGCTGTGCAC 670
Qy 698 GGGCCAAACCCAGCCAGCTCGCTCAGACGTGTGCACCGGTGTGCACCGGGTGGCACCACCTA 757
Db 671 AGACAGCAGCCAGGCCAGCTCAAGGACAAGCTCCTCTGTTGGTGGGCGCACCATCCA 730
Qy 758 TGGACTCACGCCCTCGAGCAGGGCGGCTCGCAGCAGCCACCATGATGAGCCCGTGGAGGC 817
Db 731 TGCCTTGCATGTCTGGAGAGTGGGGCTTCCGCTCCCTGCTCTCATCAACGCTGTGGAGGC 790
Qy 818 TGCACCTTGGCGGCCCAAGGAGCT 841
Db 791 CTCCTGCATCCGCACAGGGAGCT 814
US-10-161-418A-5
; Sequence 5, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-161-418A-5
Query Match 16.0%; Score 214.4; DB 14; Length 1715;
Best Local Similarity 58.0%; Pred. No. 2.7e-48;
Matches 466; Conservative 0; Mismatches 316; Indels 22; Gaps 4;
Qy 56 GCGCGTGGCTTCGTGGGCGCGGCGCATGCGGGGGCCATCGCGCAGGCGCTCATCAG 115
Db 281 GAGCGTGGCTTCATCGCGGCTGGCCAGCTGGCTTTTGGCCCTGGCCAGGCTTTTCACAGC 340
Qy 116 AGCAGGAAAGTGGAGCTCAGCACATCTAGCCAGTGCACCAACAGACAGAGAACCTATG 175
Db 341 AGCAGGCGTCTTGGCTGCGCCACAGATAATGGCTAGCTCCC---CAGACATGGAGCTGCG 397
Qy 176 TCACCTTTCAAGCTC-----TGGGTTGCCGACCGACGCACTCCAAACAGGAGGTGCT 226
Db 398 CACAGTTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACACCCCAACAGAGGACCGGT 457
Qy 227 GCAGAGCTGCCCTGCTCGTCATCTTTGCCACCAAGCCTCATGTGCTGCCAGCTGTCTGGC 286
Db 458 GCAGCAGCATGATGTGCTCTTCTGCTGTGAAGCCACACATCATCCCTTCATCCTGGA 517
Qy 287 AGAGTGGCTCCTGTGGTGCACCATGAAACACATCTTGGTGTCCGCTGGCTGTGGGGGTGC 346
Db 518 TGAATAGGCGCGACATTGAGGACAGACATTTGGTGTCTGCTGGCGCGCGGCTGCAC 577
Qy 347 TCTGAGCACCCCTGGAGGAGCTGCTG-----CCCCAAACACACGGGTGCTGGGGT 397
Db 578 CATCAGCTCCATTGAGAAGAGCTGTACAGCTTTTCCGCGCAGCCCCCAGGGTCAATCCGCTG 637
Qy 398 CTGGCCCAACTGCCCTGTGTGCTCAGGAAGGGCCATAGTATGGCGGGCGGCCGCCA 457
Db 638 CATGACCAACACTCCAGTCTGTGCGGGAGGGGCCACCGGTATGCCACAGGACGAC 697
Qy 458 CGTGGGAGCAGCAGACCAAGCTCCTGCAGCATCTGCTGGAGGCTGTGGGGCGGTGTA 517
Db 698 GCCCAGGTGGAG-GACGGGAGGCTCATGGAGCAGCTGTGAGCAGCTGGGCTTCTGCAC 756
Qy 518 GGAGTGCCTGAAGCCTAGCTGACATCCACACTGCCTCAGTGGCAGTGGGCTGGGCTT 577
Db 518 GAGTGCCTGAAGCCTAGCTGACATCCACACTGCCTCAGTGGCAGTGGGCTGGGCTT 577
```

Db 757 GGAGGTGAAGAGAGACCTGATTGATGCCGTCTCAGTGGGCTCAGTGGCAGCGGCCCGCCTTA 816  
QY 578 CGTGTGTGCAATTCCTCCAGGCGCTGGCTGAAGGAGCGGTCAAGATGGCATGCCAGCAG 637  
Db 817 CGCATTACAGCCCTGATGCCCTGGCTGATGGGGTGTGAAGATGGACTTCCAAGGCG 876  
QY 638 CTTGGGCCACCGCATCGCTCCCGAGACCCCTGCTGGGAGCGGCCAAGATGCTGTGCACGA 697  
Db 877 CTTGGCAGTCCGCTCGGGGCCCGCAGGCCCTCTCTGGGGCTGGCCAAAGATGCTGTGCAC 936  
QY 698 GGCCCAACACCCAGCCAGCTGGCTCAGACGTGTGCACCCCGGGTGGCACACCATCTA 757  
Db 937 AGAACAGACCCAGGCGCAGCTCAAGGACAAACGTCAGCTCTCTGGTGGGCGCCACCATCA 996  
QY 758 TGGACTCCAGCCCTGGAGCAGGCGGGCTGCGCAGCAGCACCATCAGCGCGCTGGAGGC 817  
Db 997 TGCCTTCATGTCTGGAGAGTGGGGCTTCGCTCCCTGCTCATCAACGCTGTGGAGGC 1056  
QY 818 TGCACCTGCCGGCCCAAGAGCT 841  
Db 1057 CTCTGCTATCCGCACACGCGAGCT 1080

## RESULT 8

US-10-161-418A-4  
; Sequence 4, Application US/10161418A  
; Publication No. US20030036078A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-055  
; CURRENT APPLICATION NUMBER: US/10/161,418A  
; CURRENT FILING DATE: 2002-09-11  
; PRIOR APPLICATION NUMBER: US 60/296,080  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,509  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1028  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (999)..(999)  
; OTHER INFORMATION: "n" is A, C, G, or T  
US-10-161-418A-4

Query Match 15.8%; Score 211.2; DB 14; Length 1028;  
Best Local Similarity 58.2%; Pred. No. 1.9e-47;  
Matches 438; Conservative 0; Mismatches 293; Indels 21; Gaps 3;  
QY 56 GCGCGTGGCTTCGTGGGCGCGCGCGCATGCGCGGGGCGCATCGCGAGGCGCTCATCAG 115  
Db 61 GAGCGTGGCTTCATCGGGCGCTGGCAGCTGCTTTGCGCGCAAGGCGTTCACAGC 120  
QY 116 AGCAGGAAGTGGAGCTCAGACATATCTAGCGCGAGTGACCAACAGACAGAACTATG 175  
Db 121 AGCAGCGCTTTGGTGGCGCCCAAGATAATGCTAGCTCC---CAGACATGGACCTGGC 177  
QY 176 TCACCTTTCAGCTC-----TGGGTTGCGGACACGACGACTCCCAACAGGAGGTGCT 226  
Db 178 CACAGTTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACCCCAACCAAGAGAGCGGT 237  
QY 227 GCAGAGCTGCCGTGCTGCTATCTTTGCCACCAAGCCTCATGTGTGCCAGCTGTCTGGC 286  
Db 238 GCAGCAGAGTATGCTGCTCTTCTGCTGTGAAGCCACACATATCCCTTCATCCCTGGA 297  
QY 287 ACAGTGGCTCCTGTGGTCCACCACTGAACATCTTGTGTCCGTGGTGTGCTGGGGTGTG 346  
Db 298 TGAATAGCGCCGACATTTGAGGACAGACATATGTTGTTGCTCTGCGCGCGCGCGTCA 357

QY 347 TCTGAGCACCTGGAGGAGCTGCTG-----CCCCCAACACACGCGTGTCTGCGGGT 397  
Db 358 CATCAGCTCCATTGAGAAGAAGCTGTACGCGTTTGGCCAGCCCGCCAGGGTTCATCCGCTG 417  
QY 398 CTTGCCCAACCTGCCCTGTGTGGTCCAGGAAGGCGCCATAGTGTATGGCGGGGCGCGCA 457  
Db 418 CATGACCAACACTCCAGTCGTGTGGGAGGAGGCGCCACCGTGTATGCCACAGCAGCA 477  
QY 458 CGTGGGAGCAGCAGACCAAGCTCTCTCAGCATCTGCTGAGAGCCTGTGTGGGCGGTGTGA 517  
Db 478 CCCCCAGTGGAGGACGGGAGGCTCATGGAGCAGTCTGTGAGCAGCGTGGGCTTCTGCAC 537  
QY 518 GGAGTGTCTGAAGCTACGTCTGACATCCACATCGGCTCAGTGGCTCAGTGGCTGGCGCTT 577  
Db 538 GGAGTGAAGAGNACCTGATGATGCGCTCAGGGGCTCAGTGGCAGCGGCCCGCCTA 597  
QY 578 CGTGTGTGCAATTCCTCCAGGCGCTGGCTGAAGAGCGGTCAAGATGGCGCATGCCAGCAG 637  
Db 598 CGCATTCACAGCCCTGGATGCCCTGGCTGATGGGGTGTGAAGATGGGACTTCCAAGCG 657  
QY 638 CTTGGCCACCGCATCGCTGCCAGACCCCTGCTGGGACGCGCAAGATGTCTGTGCACGA 697  
Db 658 CTTGGCAGTCCGCTCGGGGCGCCAGGCGCTCTCTGGGGCTGCCAAGATGTCTGTGCACT 717  
QY 698 GGGCAACACCCAGCGCGCTGCGCTCAGACGTGTGCACCCCGGGTGGCACCACCATCTA 757  
Db 718 AGAACAGCACCAGCGCAGCTCAAGGACAGGTCAAGTCTCTCTGTTGGGCGCACCATCCA 777  
QY 758 TGGACTTCAGCCCTGGAGCAGGCGCGGCTGCG 789  
Db 778 TGCCTTGATGTGCTGGAGAAGTGGGGCTTC 809

## RESULT 9

US-09-925-300-278  
; Sequence 278, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 278  
; LENGTH: 1478  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-300-278

Query Match 15.2%; Score 203.4; DB 10; Length 1478;  
Best Local Similarity 57.9%; Pred. No. 2.5e-45;  
Matches 466; Conservative 0; Mismatches 316; Indels 23; Gaps 5;

QY 56 GCGCGTGGGCTTCGTGGGCGCGCGCATGCGGGGGGCGCATCGCGAGGCGCTCATCAG 115  
Db 78 GAGCGTGGGCTTCATCGGCGCTGGCCAGCTGGCTTTTGGCCCTGGCCAAAGGCTTGCACAG 137  
QY 116 -AGCAGAAAAGTGGAGCTCAGACATAGTGGCAGTGCACCAACAGACAGAACTAT 174  
Db 138 CAGCAGCGCTTGGCTGCCACAAAGATATGCTAGCTCCC---CAGACATGGACCTGG 194  
QY 175 GTCACCTTCAAGCTC-----TGGGTTGCGGACCAACGCACTCCCAACAGGAGGTGC 225  
Db 195 CCACAGTTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACACCCCAACAAAGAGACGG 254  
QY 226 TGCAGAGCTGCCTGCTGCTCATCTTTGCCAACCAAGCTCATGTGCTGCCAGCTGTCTCTCG 285



Db	720	GGGCGTGCCTATGCAATTCATGCGCTGAGGCAATTGGCTGATGTTGGGTGAGCAATGGGT	7719
QY	627	ATGCCACGACGCCTGGCCCCACCGCATGCTGCCAGACCCCTCTGGGACGCGCCAAGATG	686
Db	780	TTGCCACGCGCCTGGCAATTCACAACTCGGGGCCAGGCTTTGCTGGGAGCTGCCAAGATG	833
QY	687	CTGTGTGCACGAGGGCCCAACACCCAGCCAGCTCCGCTCAGACGTCTGCACCCCGGTGGC	746
Db	840	CTGCTGGACTCGGAGCAGCATCTGCACGCTTAAGGACAATGTCTGCTCCCTCTGGGGGA	899
QY	747	ACCCACCATCTATGACTCCACGCCCTGGAGCAGGGCGGCTCGAGCAGCCACCATGAGC	806
Db	900	GCCACCATCCAGCCCTGCACCTTCTAGAGAGTGGGGCTTCCGCTCTCTGCTCATCAAT	959
QY	807	GCGGTGAGGCTGCCACCTGCGGGGCCAAGGAGCT	841
Db	960	GCAGTTGAGGCTCTCTGTATCCGAACACGAGAGCT	994

RESULT 12  
US-10-161-418A-9  
; Sequence 9, Application US/10161418A  
; Publication No. US20030036078A1

```

PUBLICATION NO. 0820030670/SH
;
; GENERAL INFORMATION:
;
; APPLICANT: EXELIXIS, INC.
;
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
;
; FILE REFERENCE: EX02-055
;
; CURRENT APPLICATION NUMBER: US/10/161,418A
;
; CURRENT FILING DATE: 2002-09-11
;
; PRIOR APPLICATION NUMBER: US 60/296,080
;
; PRIOR FILING DATE: 2001-06-05
;
; PRIOR APPLICATION NUMBER: US 60/328,509
;
; PRIOR FILING DATE: 2001-10-10
;
; NUMBER OF SEQ ID NOS: 14
;
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 9
;
; LENGTH: 999
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-10-161-418A-9

```

Query Match	12.58	Score 167	DB 14	Length 999
Best Local Similarity	56.08	Pred. No. 1.6e-35		
Matches 362	Conservative	0	Mismatches 275	Indels 10
				Gaps 2

[illegible]

Query Match	15.1k	Score	201.8	DB	9	Length	1742
Best Local Similarity	55.0k	Prd.	No.	6.9e-45			
Matches	448	Conservative	0	Mismatches	352	Indels	15
	Gaps	2					
QY	42	GAGCGCTCTCGGGCGCGTGGGCTTCGTCGGCGCGGGCCGCATGGGCGGGCCATCGCG	101				
Db	180	GAACCGCGACATGAGCGTGGGCTTCATCGGGCGCGCCAGCTGGCCTATGCTCTGGCG	239				
QY	102	CAGGCGCTCTATCAGAGAGAAAGTGGAGCTCAGCACATCTAGCGAGTGCACCAACA	161				
Db	240	CGGGGCTTCACGGCCGAGCATCTCTCGGCTCAAGAATAATAGCAGCTCCCCAGAA	299				
QY	162	GACAGGAAC-----TATGTCACTTTCAGCTCTGGGTTCGGGACCGACGACATCCCAAC	215				
Db	300	ATGAACCTGCCACGGTGTCCGGCTCAGGAAGATGGGTGTGAACCTGCACAGCAGCAAC	359				
QY	216	CAGGAGTGTGCAGAGTGCCTGCTGCTCATCTTTGCCACAAAGCCTCATGTGCTCCCA	275				
Db	360	AAGGAGAGCGTGAAGCACAGCAGCTCTGTTCTTGGGTGTGAAGCCACATATCATCCCC	419				
QY	276	GCTGCTCTGCACAGGTGGCTCTGTGGTCAACCTGAACACATCTTGGTGTCCGTGGCT	335				
Db	420	TTCACTCTGATGAGATTGGGGCGACGTGCAGCCAGACACATCTGTGTTCTCTGTGGC	479				
QY	336	GCTGGGTGTCTCTGACACACCTGGAGGAGCTGCTG-----CCCCAAACACACGG	386				
Db	480	GCTGCTGTCAACATCAGCTCTGTGGAGAAGAAGCTGATGGCATCCAGCCAGCCCCAAA	539				
QY	387	GTGCTGGCGCTCTTGCCCAACCTGCTGTGGTTCAGGAAGGGGCCATGATGATGGCG	446				
Db	540	GTGATTCGCTCATGACACACACACTGTGGTGTAGTCAGGAAGGCGCTACAGTGTAGCC	599				
QY	447	CGGGGCCGCCACCTGGGGAGCAGCAGACCAAGCTTCCTGCAGCATCTGCTGGAGGCGCTG	506				
Db	600	ACGGGCACCCATCCCTGTGTGGAGGATGGCAGCTCCTGGAGCAGCTCATGAGCAGCGTG	659				
QY	507	GGGCGGTGTGAGGAGTGGCTGAGCCCTACGTGCATCCACATCGCCTCAGTGGCAGT	566				
Db	660	GGCTTCTGCACTCAGGTGGAAAGAGACCTTCATCGATGCCGTACGGGGCTCAGTGGCAGC	719				
QY	567	GGGCTGGCCCTTCGTGTGTCATTCTCCAGGGCCCTGGCTGAAGGAGCGCTCAAGATGGC	626				





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 396 Seconds  
(without alignments)  
9120.829 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcgagctgtctcga.....ggggcagcagcatatggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_19Jun03:*			
1:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*						
2:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*						
3:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*						
4:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*						
5:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*						
6:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*						
7:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*						
8:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*						
9:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*						
10:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*						
11:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*						
12:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*						
13:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*						
14:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*						
15:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*						
16:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*						
17:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*						
18:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*						
19:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*						
20:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*						
21:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*						
22:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*						
23:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*						
24:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*						
25:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1338	100.0	1338	21	Human DNA encoding
2	1307	97.7	2331	22	Human cDNA sequenc
3	1307	97.7	2331	25	P5CR related nucle
4	1302.2	97.3	2496	22	Human dihydropyrro
5	1291.8	96.5	2615	22	Human cDNA sequenc
6	1127.8	84.3	1178	25	P5CR related nucle
7	857	64.1	860	25	P5CR related nucle
8	737.6	55.1	794	22	Human cDNA clone (

9	723	54.0	5055	22	AAK51999	Human polynucleoti
10	599.4	44.8	5647	22	AA542169	Genomic sequence #
11	486.6	36.4	563	22	AAH06264	Human cDNA clone (
12	241.8	18.1	270	22	AA541342	cDNA encoding nove
13	226.4	16.9	1848	25	ABA00760	P5CR related nucle
14	224.8	16.8	1792	24	ABT10919	Human breast cance
15	224.8	16.8	1792	24	ABN95903	Gene #2401 used to
16	224.8	16.8	1792	25	ABA00758	P5CR related nucle
17	224.8	16.8	1889	24	ABO54355	Human ovarian anti
18	214.4	16.0	1715	25	ABA00762	P5CR related nucle
19	211.2	15.8	1028	25	ABA00761	P5CR related nucle
20	203.4	15.2	1478	21	AAF15843	Human prostate can
21	201.8	15.1	1708	25	ABA00765	P5CR related nucle
22	201.8	15.1	1742	21	AAA88073	Human delta 1-pyrr
23	201.8	15.1	1742	22	AAD20894	Human delta 1-pyrr
24	201.8	15.1	1742	24	ABK87030	Human Py-CR encodi
25	197	14.7	1121	22	AAF81847	cDNA encoding nove
26	194.4	14.5	251	22	AA541363	Human dithp enzyme
27	183.6	13.7	1587	25	ACA46089	P5CR related nucle
28	167	12.5	999	25	ABA00766	Human cDNA sequenc
29	148.8	11.1	3149	22	AAH14378	Drosophila melanog
30	143	10.7	867	23	ABL10077	Drosophila melanog
31	127.6	9.5	3120	23	ABL10076	Drosophila melanog
32	127.6	9.5	4978	23	ABL10046	Drosophila melanog
33	118.2	8.8	1183	23	ABL29275	Drosophila melanog
34	112.4	8.4	1868	23	AAS84881	DNA encoding novel
35	110.2	8.2	3244	23	ABL29274	Drosophila melanog
36	110.2	8.2	3284	23	ABL29246	Drosophila melanog
37	105	7.8	1163020	24	ABQ67197	Listeria innocua c
38	105	7.8	3011208	24	ABQ69245	Listeria innocua c
39	94.2	7.0	1837	22	AAH48471	Escherichia coli p
40	87.4	6.5	56050	23	AAS59849	Propionibacterium
41	84.4	6.3	810	22	AAH65425	C glutamicum codin
42	84.4	6.3	933	22	AAF71901	Corynebacterium gl
43	84.4	6.3	349980	22	AAH68525	C glutamicum codin
44	84.2	6.3	2944528	24	ABA03041	Listeria monocytog
45	82.6	6.2	6743	24	ABQ71038	Listeria monocytog

ALIGNMENTS

RESULT 1

AAA09388

ID AAA09388 standard; DNA; 1338 BP.

XX AAA09388;

AC AAA09388;

XX 10-AUG-2000 (first entry)

DT 10-AUG-2000 (first entry)

DE Human DNA encoding OXRE-14.

XX OXRE-14; oxidoreductase; pyrroline-5-carboxylate-reductase;

KW antiproliferative; anticancer; hepatotropic; antiviral; antiasthmatic;

KW neurotropic; neuroprotectant; antiparkinsonian's; antisclerotic;

KW anxiolytic; antischizophrenic; anti-inflammatory; antiarthritic; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 30..854

FT FT /\*tag= a

FT FT /product= OXRE-14

XX WO200020604-A2.

XX 13-APR-2000.

XX 06-OCT-1999; 99WO-US23434.

XX 06-OCT-1998; 98US-0172227.

XX 02-DEC-1998; 98US-0155202.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.

XX 10-MAR-1999; 99US-0123911.



XX (INCY-) INCYTE PHARM INC.  
XX Lal P, Guegler KJ, Gorgone GA, Corley NC, Baughn MR, Tang YN;  
PI Hillman JL, Bandman O, Azimzai Y, Au-Young J, Yue H, Lu DAM;  
PI Yang J;  
XX WPI; 2000-303785/26.  
DR P-PSDB; AAY92517.  
XX Purified polypeptide for treating or preventing disorders associated  
PT with decreased expression or activity of oxidoreductase molecules  
XX  
PS Claim 9; Page 96; 97pp; English.  
XX AAA09375-89 encode oxidoreductases, designated OXRE-1 to -15. OXRE-14  
CC has identity with pyrroline-5-carboxylate-reductase.  
CC The polypeptides are useful for treating or preventing a disorder  
CC associated with decreased expression or activity of OXRE. Antagonists  
CC of OXRE are useful for treating or preventing a disorder associated with  
CC increased expression or activity of OXRE. The disorders include cell  
CC proliferative disorders (cirrhosis, hepatitis), cancer (leukemia,  
CC melanoma), hypopituitarism and hyperpituitarism, hypothyroidism and  
CC hyperthyroidism, metabolic disorders (Addison's disease, cystic  
CC fibrosis), reproductive disorders (infertility, ovulatory defects),  
CC neurological disorders (Alzheimer's disease, Parkinson's disease,  
CC multiple sclerosis), mental disorders (anxiety, schizophrenia),  
CC autoimmune/inflammatory disorders (acquired immunodeficiency syndrome  
CC (AIDS), asthma, osteoarthritis), and viral infections. The  
CC polynucleotides may be used in Southern or Northern analysis, polymerase  
CC chain reaction (PCR), or in enzyme-linked immunosorbent assays (ELISA).  
XX  
SQ Sequence 1338 BP; 225 A; 423 C; 440 G; 250 T; 0 other;  
Query Match 100.0%; Score 1338; DB 21; Length 1338;  
Best Local Similarity 100.0%; Pred. No. 3 5e-289;  
Matches 1338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTGAGCGCAGTGTGTCGAGGCAACAAAGATGCGAGTGGGAGCGCTTCCGCGCGCG 60  
Db 1 GGTGAGCGCAGTGTGTCGAGGCAACAAAGATGCGAGTGGGAGCGCTTCCGCGCGCG 60  
QY 61 TGGGCTTCGTTGGGCGCGCGCATGCGGGGGCCATCGCGAGGCGCTCATCAGAGCAG 120  
Db 61 TGGGCTTCGTTGGGCGCGCGCGCATGCGGGGGCCATCGCGAGGCGCTCATCAGAGCAG 120  
QY 121 GAAAGTGGAAGCTCAGCACAATCTGCGCAGTGCACCAACAGACAGGAACCTATGTCACT 180  
Db 121 GAAAGTGGAAGCTCAGCACAATCTGCGCAGTGCACCAACAGACAGGAACCTATGTCACT 180  
QY 181 TTCAAGCTCTGGGTTGCCGACACCGACACTTCCAAACAGGAGTGTCTGCAGAGTGCCTGC 240  
Db 181 TTCAAGCTCTGGGTTGCCGACACCGACACTTCCAAACAGGAGTGTCTGCAGAGTGCCTGC 240  
QY 241 TCGTCATCTTGGCCACCAACGCTCATGTGCTGCCAGCTGTCTGGCAGAGTGGCTCCG 300  
Db 241 TCGTCATCTTGGCCACCAACGCTCATGTGCTGCCAGCTGTCTGGCAGAGTGGCTCCG 300  
QY 301 TGGTCACTCTGAACACATCTTGGTGTCCGTGGGCTGTCTGTGAGCACCGCTGG 360  
Db 301 TGGTCACTCTGAACACATCTTGGTGTCCGTGGGCTGTCTGTGAGCACCGCTGG 360  
QY 361 AGGAGCTGTGCCCCCAACACACACGCGGTGTCTGGGGTCTTGGCCAACTGCCCTGTGTGG 420  
Db 361 AGGAGCTGTGCCCCCAACACACACGCGGTGTCTGGGGTCTTGGCCAACTGCCCTGTGTGG 420  
QY 421 TCCAGGAAGGGCCATAGTATGCGCGGGGGCCCGCAGCTGGGGAGCAGCAGACCAAC 480  
Db 421 TCCAGGAAGGGCCATAGTATGCGCGGGGGCCCGCAGCTGGGGAGCAGCAGACCAAC 480  
QY 481 TCCTGCAGCATCTGCTGGAGCGCTGTGGGCGGTGTGAGGAGTGTCTGAAGCCTACGTCG 540  
Db 481 TCCTGCAGCATCTGCTGGAGCGCTGTGGGCGGTGTGAGGAGTGTCTGAAGCCTACGTCG 540

QY 541 ACATCCACACTGGCCTCAGTGGCAGTGGCGTGGCCTTCGTGTGTGTCATCTCCGAGGCC 600  
Db 541 ACATCCACACTGGCCTCAGTGGCAGTGGCGTGGCCTTCGTGTGTGTCATCTCCGAGGCC 600  
QY 601 TGGCTGAAGGAGCCCTCAAGATGGCATGCCAGAGCAGCTGGCCACCGCATCGCTGCC 660  
Db 601 TGGCTGAAGGAGCCCTCAAGATGGCATGCCAGAGCAGCTGGCCACCGCATCGCTGCC 660  
QY 661 AGACCCCTGCTGGGAGCGCCAGATGCTGTCAGAGGAGCCACACACCCAGCCAGCTGC 720  
Db 661 AGACCCCTGCTGGGAGCGCCAGATGCTGTCAGAGGAGCCACACACCCAGCCAGCTGC 720  
QY 721 GCTCAGAGCTGTGCACCCCGGCTGGCACACCATATATGGACTCCACGCCCTGGAGCAGG 780  
Db 721 GCTCAGAGCTGTGCACCCCGGCTGGCACACCATATATGGACTCCACGCCCTGGAGCAGG 780  
QY 781 GCGGCTCGGAGCAGCCACCATGAGCGCGTGGAGGCTGCCACCTCCGCGGCGCAAGAGC 840  
Db 781 GCGGCTCGGAGCAGCCACCATGAGCGCGTGGAGGCTGCCACCTCCGCGGCGCAAGAGC 840  
QY 841 TCAGCAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCCCTGCTGTGTGCCCTGCTCTCC 900  
Db 841 TCAGCAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCCCTGCTGTGTGCCCTGCTCTCC 900  
QY 901 CTGTCTCCCTTCCCTGAGGACTGCGGCTCCCTCCCTGCTGCATGAGGCTTCTCTACTGC 960  
Db 901 CTGTCTCCCTTCCCTGAGGACTGCGGCTCCCTCCCTGCTGCATGAGGCTTCTCTACTGC 960  
QY 961 TCCTTCTCCCTTGCACAGGAAATGCGAGGCGGAGGACTTGGGAGGTTCCAGCAGCGG 1020  
Db 961 TCCTTCTCCCTTGCACAGGAAATGCGAGGCGGAGGACTTGGGAGGTTCCAGCAGCGG 1020  
QY 1021 GGGAGCCCCGACAGTGGGACACTTCCCTCCCTCCCGAGTGCAGAGGACCGTGGTGG 1080  
Db 1021 GGGAGCCCCGACAGTGGGACACTTCCCTCCCTCCCGAGTGCAGAGGACCGTGGTGG 1080  
QY 1081 TGGCTCTGCCCTTCTGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1140  
Db 1081 TGGCTCTGCCCTTCTGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1140  
QY 1141 AGAGATGGGCTTGGTTCATTTGCCCGCATGGTTGGGAGTGGTGGGAGGTTCCAGCAG 1200  
Db 1141 AGAGATGGGCTTGGTTCATTTGCCCGCATGGTTGGGAGTGGTGGGAGGTTCCAGCAG 1200  
QY 1201 AACTTACGCTAACAGGCACGCTGCCCAATGCCCTGGTCTGGAGCTGGAGCTTGCCTTTG 1260  
Db 1201 AACTTACGCTAACAGGCACGCTGCCCAATGCCCTGGTCTGGAGCTGGAGCTTGCCTTTG 1260  
QY 1261 GCTTTTCCAAAGTGGGCTCGTGCAGCTTACAGCCAGGCGCGCTGCCTCATCTCAGCTCTAGG 1320  
Db 1261 GCTTTTCCAAAGTGGGCTCGTGCAGCTTACAGCCAGGCGCGCTGCCTCATCTCAGCTCTAGG 1320  
QY 1321 GGCAGGACATATGGGT 1338  
Db 1321 GGCAGGACATATGGGT 1338  
RESULT 2  
AAH18288  
ID AAH18288 standard; cDNA; 2331 BP.  
XX  
XX AAH18288;  
XX AC  
XX DT 26-JUN-2001 (first entry)  
XX  
XX Human cDNA sequence SEQ ID NO:18268.  
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
KW Homo sapiens.  
XX  
XX EP1074617-A2.  
PN

XX PD. 07-FEB-2001.  
XX PF. 28-JUL-2000; 2000EP-0116126.  
XX PR. 29-JUL-1999; 99JP-0248036.  
XX PR. 27-AUG-1999; 99JP-0300253.  
XX PR. 11-JAN-2000; 2000JP-0118776.  
XX PR. 02-MAY-2000; 2000JP-0183767.  
XX PR. 09-JUN-2000; 2000JP-0241899.  
XX PA. (HELI-) HELIX RES INST.  
XX PI. Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX PT. Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX PS. Claim 8; SEQ ID 18268; 2537pp + CD ROM; English.  
XX CC. The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX SQ. Sequence 2331 BP; 398 A; 753 C; 748 G; 432 T; 0 other;  
Query Match 97.7%; Score 1307; DB 22; Length 2331;  
Best Local Similarity 99.2%; Pred. No. 3.3e-282;  
Matches 1313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 15 GTCGAGGCAACAAGATGGCAGCTGCGGAGCCGCTCCCGCGCGCTGCGCTTCGTGGGC 74  
Db 12 GTCGAGGCAACAAGATGGCAGCTGCGGAGCCGCTCCCGCGCGCTGCGCTTCGTGGGC 71  
Qy 75 GCGGGCGGATGCGGGGGGCATCGCGCAGGGGCTCATCAGAGCAGGAAAGTGGAGCT 134  
Db 72 GCGGGCGGATGCGGGGGGCATCGCGCAGGGGCTCATCAGAGCAGGAAAGTGGAGCT 131  
Qy 135 CAGCACAATAGTGGCAGTGCACCAACAGACAGAACCTATGTCACCTTTCAAGCTCTGGGT 194  
Db 132 CAGCACAATAGTGGCAGTGCACCAACAGACAGAACCTATGTCACCTTTCAAGCTCTGGGT 191  
Qy 195 TGGCGGACCAACGACATCCCAACAGGAGGTGCTGCAGAGCTGGCTGCTCATCTTTGGC 254  
Db 192 TGGCGGACCAACGACATCCCAACAGGAGGTGCTGCAGAGCTGGCTGCTCATCTTTGGC 251  
Qy 255 ACCAAGCCTCATGTGCTGCCAGCTGTCTCGGACAGAGGTGGCTGCTGTGCTACCACTGAA 314  
Db 252 ACCAAGCCTCATGTGCTGCCAGCTGTCTCGGACAGAGGTGGCTGCTGTGCTACCACTGAA 311

Qy 315 CACATCTTGGTGTCCGTGGCTGGGTGTCTCTGAGCACCCTGGAGGAGCTGCTGCC 374  
Db 312 CACATCTTGGTGTCCGTGGCTGGGTGTCTCTGAGCACCCTGGAGGAGCTGCTGCC 371  
Qy 375 CCAAAACACACGGGTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGTGTCAGAGGAGGCC 434  
Db 372 CCAAAACACACGGGTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGTGTCAGAGGAGGCC 431  
Qy 435 ATAGTGATGGCGCGGGGCGCCACGTGGGAGCAGCAGACCAAGCTCTCTGACGATCTG 494  
Db 432 ATAGTGATGGCGCGGGGCGCCACGTGGGAGCAGCAGACCAAGCTCTCTGACGATCTG 491  
Qy 495 CTGAGAGCCTGTGGCGGTGTGAGAGGTGCTTGAAGCTACGTCGACATCCACACATGGC 554  
Db 492 CTGAGAGCCTGTGGCGGTGTGAGAGGTGCTTGAAGCTACGTCGACATCCACACATGGC 551  
Qy 555 CTCAGTGGCAGTGGCGGTGCTTGTGTGTGTCATTTCCGAGGCCCTGGCTGAAGAGGCC 614  
Db 552 CTCAGTGGCAGTGGCGGTGCTTGTGTGTGTCATTTCCGAGGCCCTGGCTGAAGAGGCC 611  
Qy 615 GTCAGATGGCAGTGGCGGAGCCTGCGCCACCGCATCGCTGCGCAGACCTGCTGGGG 674  
Db 612 GTCAGATGGCAGTGGCGGAGCCTGCGCCACCGCATCGCTGCGCAGACCTGCTGGGG 671  
Qy 675 ACGGCCAAGATGCTGTCAGAGGCGCAACACCCAGCCAGCTGCGCTCAGACGTGTGC 734  
Db 672 ACGGCCAAGATGCTGTCAGAGGCGCAACACCCAGCTGCGCTCAGACGTGTGC 731  
Qy 735 ACCCGGGTGGCAGCACCACATCTATGGACTCCACGCCCTGGAGCAGGCGGCTCGCAGCA 794  
Db 732 ACCCGGGTGGCAGCACCACATCTATGGACTCCACGCCCTGGAGCAGGCGGCTCGCAGCA 791  
Qy 795 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCGGGGCCAAGAGAGCTCAGCAGAAAGTAG 854  
Db 792 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCGGGGCCAAGAGAGCTCAGCAGAAAGTAG 851  
Qy 855 GCTGGGCTCTGGCCATCTCTTCTGCGCTCTGCGCCCTGCTCTCTCTGCTGCTCTTCCC 914  
Db 852 GCTGGGCTCTGGCCATCTCTTCTGCGCTCTGCGCCCTGCTCTCTCTGCTGCTCTTCCC 911  
Qy 915 CTGAGGACTGCGGCTCCCTCCTCTGTCATGAGGCTCTCTCTCTCTCTCTCTCTCTCT 974  
Db 912 CTGAGGACTGCGGCTCCCTCCTCTGTCATGAGGCTCTCTCTCTCTCTCTCTCTCTCT 971  
Qy 975 CACAGGAAATGACAGGCGGAGGCTTGGAGGTTCCAGCAGCGGGGAGGCCCGGACCA 1034  
Db 972 CACAGGAAATGACAGGCGGAGGCTTGGAGGTTCCAGCAGCGGGGAGGCCCGGACCA 1031  
Qy 1035 GTGGGGACATCT 1094  
Db 1032 GTGGGGACATCT 1091  
Qy 1095 GCTCAGTGAAGCCACCTTGTGCAACATTTGGTCTGAGGGGCCCAAGAGATGGGCTCT 1154  
Db 1092 GCTCAGTGAAGCCACCTTGTGCAACATTTGGTCTGAGGGGCCCAAGAGATGGGCTCT 1151  
Qy 1155 GGTCAATTTGCCGATGCTTGGGAGTTGGGAGTTGGAGCTTGGCTTTGGCTTTTCCAGTGG 1214  
Db 1152 GGTCAATTTGCCGATGCTTGGGAGTTGGGAGTTGGAGCTTGGCTTTTCCAGTGG 1211  
Qy 1215 GGCAGGCTGGCCCAATGCTGCTGGAGCTTGGAGCTTGGCTTTGGCTTTTCCAGTGG 1274  
Db 1212 GGCAGGCTGGCCCAATGCTGCTGGAGCTTGGAGCTTGGCTTTTCCAGTGG 1271  
Qy 1275 CTCTGTCAGCTACAGCCAGGCGGTGCTCATCTCTAGCTCTAGGGGCGCAGGACATAG 1334  
Db 1272 TCCGTGTCAGCTACAGCCAGGCGGTGCTCATCTCTAGCTCTAGGGGCGCAGGACATAT 1331  
Qy 1335 GGG 1337  
Db 1332 GGG 1334



	QY	1335 GGG 1337 	
	Dd	1332 GGG 1334	
<b>RESULT 4</b>			
	AAH77597		
ID	AAH77597 standard; cDNA; 2496 BP.		
XX	AC AAH77597;		
XX	22-OCT-2001 (first entry)		
DT	XX		
DE	Human dihydropyrrrole-5-carboxylate reductase 30 cdNA.		
KW	Human; dihydropyrrrole-5-carboxylate reductase 30; cancer; cytostatic;		
KW	human immunodeficiency virus; HIV; infection; immunological disease;		
KW	Inflammatory disease; ss.		
XX	OS Homo sapiens.		
XX	CN1298002-A.		
PX	PN XX		
PD	XX 06-JUN-2001.		
PF	XX 24-NOV-1999; 99CN-0124090.		
PR	XX 24-NOV-1999; 99CN-0124090.		
PA	(SHAN-) SHANGHAI BORONG GENE DEV CO LTD.		
PI	Mao Y, Xie Y;		
DR	WI: 2001-489680/54.		
DR	P-PSDB; AAG66956.		
XX	Human dihydropyrrrole-5-carboxylate reductase 30 as one new kind of polypeptide and polynucleotides encoding this polypeptide -		
PT	Claim 6; Page 19-20 (disclosure); 26pp; Chinese.		
PS	The invention relates to a novel polypeptide, human		
CC	dihydropyrrrole-5-carboxylate reductase 30, polynucleotides encoding this polypeptide and a DNA recombination process to produce the polypeptide. The polypeptide is useful for treating various diseases, such as malignant tumours, nosohaemia, HIV infection, immunological diseases and inflammatory diseases. The invention also provides an antibody against the polypeptide. The present sequence encodes the polypeptide of the invention.		
CC	Sequence 2496 BP; 451 A; 793 C; 789 G; 463 T; 0 other;		
SQ	Query Match 97.3%; Score 1302.2; DB 22; Length 2496; Best Local Similarity 99.0%; Pred. No. 3.9e-281; Matches 1310; Conservative 0; Mismatches 13; Indels 0; Gaps 0;		
QY	15 GTCCGAGGCCAACAAGATGCAGTCGGAGCGTCTCCGCGCGGTGGGCTTCGTGGGC 74		
Dd	14 GTCCGAGGCCAACAGATGCAGTCGGAGCGTCTCCGCGCGGTGGGCTTCGTGGGC 73		
QY	75 GC GGCGCGCATGCGGGGGCCATCGCGCAGGCGCTCATCAGACGAGGAAGAAGCT 134		
Dd	74 GC GGCGCGCATGCGGGGGCCATCGCGCAGGCGCTCATCAGACGAGGAAGAAGCT 133		
QY	135 CAGCACATACTGCCAGTGCCACCACACAGACGAAACCTATGTCATTCAAGCTCTGGGT 194		
Dd	134 CAGCACATACTGCCAGTGCCACCACACAGACGAAACCTATGTCATTCAAGCTCTGGGT 193		
QY	195 TGCCGGACACGACACTCCCAACAGAGGTGCTGCAGAGCTGCCTGCTCATCTTTGCC 254		
Dd	194 TGCCGGACACGACACTCCCAACAGAGGTGCTGCAGAGCTGCCTGCTCATCTTTGCC 253		

Db	1334	GGG 1336		QY	27	AAGATGCGAGCTGCGAGCCCTCTCCGCGCGCTGGCTTCGTGGCGCGGGCCGCGCATG	86
				Db	2615	AAGATGCGAGCTGCGAGCCCTCTCCGCGCGCTGGCTTCGTGGCGCGGGCCGCGCATG	2556
				QY	87	GCGGGGCGCATGCGCGAGGGCTCATCAGAGCAGGAAAAGTGAAGCTCAGCACACTACTG	146
				Db	2555	GCGGGGCGCATGCGCGAGGGCTCATCAGAGCAGGAAAAGTGAAGCTCAGCACACTACTG	2496
				QY	147	GCCAGTGCACCAACAGACAGAACCTATGTCACCTTTCACCTTCAGCTCTGGGTTGCCGGACACG	206
				Db	2495	GCCAGTGCACCAACAGACAGAACCTATGTCACCTTTCACCTTCAGCTCTGGGTTGCCGGACACG	2436
				QY	207	CACCTCCAAACAGGAGTGCCTGCAGAGCTGCTGCTCATCTTTGGCCACCAAGCCTCAT	266
				Db	2435	CACCTCCAAACAGGAGTGCCTGCAGAGCTGCTGCTCATCTTTGGCCACCAAGCCTCAT	2376
				QY	267	GTGCTGCCAGCTGCTGGCAGAGTGGCTCCTGTTGGTCACTACTCAACACACTCTTGGTG	326
				Db	2375	GTGCTGCCAGCTGCTGGCAGAGTGGCTCCTGTTGGTCACTACTCAACACACTCTTGGTG	2316
				QY	327	TCCGTGCTGCTGGGGTGTCTCTGAGCACCTTGAGGAGCTGCTGCCCCCAACACACGG	386
				Db	2315	TCCGTGCTGCTGGGGTGTCTCTGAGCACCTTGAGGAGCTGCTGCCCCCAACACACGG	2256
				QY	387	GTGCTGCGGGTCTTGCCCAACCTGCTGTGTGTCCAGGAAGGGCCATAGTAGGGG	446
				Db	2255	GTGCTGCGGGTCTTGCCCAACCTGCTGTGTGTCCAGGAAGGGCCATAGTAGGGG	2196
				QY	447	CGGGGCGCCACGTGGGGAGCAGGAGCAACCTCTGCAAGCATCTGCTGGAGGCTGT	506
				Db	2195	CGGGGCGCCACGTGGGGAGCAGGAGCAACCTCTGCAAGCATCTGCTGGAGGCTGT	2136
				QY	507	GGGCGGTGTGAGGAGTGCCTGAAGCTACGTGACATCCACACTGCGCTCAGTGCAGT	566
				Db	2135	GGGCGGTGTGAGGAGTGCCTGAAGCTACGTGACATCCACACTGCGCTCAGTGCAGT	2076
				QY	567	GGCGTGCCCTTCTGTGTGCAATTCGAGGCGCTGGCTGAAGAGCGCTCAAGATGGG	626
				Db	2075	GGCGTGCCCTTCTGTGTGCAATTCGAGGCGCTGGCTGAAGAGCGCTCAAGATGGG	2016
				QY	627	ATGCCAGCAGCCTGGCCCAACCCGCTGCTGCCAGACCTGCTGGGGAGCGGCAAGATG	686
				Db	2015	ATGCCAGCAGCCTGGCCCAACCCGCTGCTGCCAGACCTGCTGGGGAGCGGCAAGATG	1956
				QY	687	CTGCTGCAGCAGGGCCCAACCCAGCTGCGGTGACAGCTGTGCACCCCGGTGGC	746
				Db	1955	CTGCTGCAGCAGGGCCCAACCCAGCTGCGGTGACAGCTGTGCACCCCGGTGGC	1896
				QY	747	ACCACCATCTATGAGCTCCACGCCCTGGAGCAGGGGGCTGGCAGCAGCCATGAGC	806
				Db	1895	ACCACCATCTATGAGCTCCACGCCCTGGAGCAGGGGGCTGGCAGCAGCCATGAGC	1836
				QY	807	GCGTGAGGCTGCCACCTGCGGGCCCAAGGAGCTCAGCAGAAAGTAGGCTGGCTCG	866
				Db	1835	GCGTGAGGCTGCCACCTGCGGGCCCAAGGAGCTCAGCAGAAAGTAGGCTGGCTCG	1776
				QY	867	CCATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	926
				Db	1775	CCATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1716
				QY	927	GCTCCCTCCCTCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	986
				Db	1715	GCTCCCTCCCTCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1656
				QY	987	CAGGGGCGAGGCTTGGAGGTTCCAGAGCGGGGGAGCCCGACAGCTGGGGACACATC	1046
				Db	1655	CAGGGGCGAGGCTTGGAGGTTCCAGAGCGGGGGAGCCCGACAGCTGGGGACACATC	1596
				QY	1047	CTCCCTCCCGAGTGCAGCAAGGACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1106
				Db	1595	CTCCCTCCCGAGTGCAGCAAGGACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1536
				QY	1107	CCACTTGTGTGAACATTTGTTCTGAGGGGGCCCAAGAGATGGGCTCTTGGTCAATTTGCC	1166

RESULT 5  
AAH14370/c  
ID AAH14370 standard; cDNA: 2615 BP.  
XX  
AC AAH14370;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA sequence SEQ ID NO:11778.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN EF1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs.  
XX  
PS Claim 8; SEQ ID 11778; 2537pp + CD ROM; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides, and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 2615 BP; 492 A; 831 C; 818 G; 474 T; 0 other;

Query Match 96.5%; Score 1291.8; DB 22; Length 2615;  
Best Local Similarity 99.1%; Pred. No. 8.3e-279;  
Matches 1299; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Db 1535 CCACCTTGGCCCAACATTTGGTTCTGAGGGGCCCAAGAGATGGCTCTTGGTCAATTTGCC 1476  
QY 1167 GCATGGTTGGGCACTTGGTTGAGGCCATGAACAGAACTTACGGTAACAGCAGCGCTGGC 1226  
Db 1475 GCATGGTTGGGCACTTGGTTGAGGCCATGAACAGAACTTACGGTAACAGCAGCGCTGGC 1416  
QY 1227 CCAATGCTGGTCTGGAGCTGGAGCTTTGCCCTTTGGCTTTTCCAAAGTGGGCTCGTGCAAGCTA 1286  
Db 1415 CCAATGCTGGTCTGGAGCTGGAGCTTTGCCCTTTGGCTTTCCAGGTGGCTCCGTGCAGCTA 1356  
QY 1287 CAGCCAGCGGCTGCCTCATCTCAGCTCTAGGGGGCAGCAGCATATFGGGG 1337  
Db 1355 CAGCCAGCGGCTGCCTCATCTCAGCTCTAGGGGGCAGCAGCATATFGGG 1305

RESULT 6

ABA00764  
ID ABA00764 standard; DNA: 1178 BP.  
XX AC ABA00764;  
XX  
DT 18-MAR-2003 (first entry)  
XX  
DE P5CR related nucleic acid #7.  
XX  
KW Pyroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;  
KW pyroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200299043-A2.  
XX  
PD 12-DEC-2002.  
XX  
PF 03-JUN-2002; 2002WO-US17319.  
XX  
PR 05-JUN-2001; 2001US-296080P.  
PR 10-OCT-2001; 2001US-328509P.  
XX  
XX (EXEL-) EXELIS INC.  
XX  
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
PI Engst S;  
XX  
DR WPI: 2003-140606/13.  
XX  
PT Identifying p53 pathway modulating agents with P5CR genes, useful for  
PT the diagnosis and treatment of disorders associated with defects in the  
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and  
PT ovary -  
XX  
PS Disclosure; Page 44-45; 53pp; English.  
XX  
CC The sequences given in ABA00758-66 represent pyrroline 5 carboxylate  
CC reductase (P5CR) related nucleic acid sequences. P5CR catalyses the  
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to  
CC proline. These sequences may be used in the method of the invention  
CC for identifying a candidate p53 pathway modulating agent. The method  
CC comprises providing an assay system comprising a purified P5CR  
CC polypeptide or nucleic acid, or a functionally active fragment or  
CC derivative, contacting the assay system with a test agent, where the  
CC system provides a reference activity, and detecting a test agent-biased  
CC activity of the assay system. The method of the present invention is  
CC useful for the diagnosis and treatment of disorders associated with  
CC defects in the p53 pathway, such as cancer of the breast, colon,  
CC kidneys, lung and ovary.  
XX  
SQ Sequence 1178 BP; 211 A; 379 C; 383 G; 205 T; 0 other;

Query Match 84.3%; Score 1127.8; DB 25; Length 1178;  
Best Local Similarity 99.8%; Pred. No. 2.8e-242;  
Matches 1129; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 GTCCGAGGCAACAAGATGCAGCTGGGAGCCGTCTCCGCGCGCGTGGGCTTCGTGGGC 74  
|||||  
Db 32 GTCCGAGGCAACAAGATGCAGCTGGGAGCCGTCTCCGCGCGCGTGGGCTTCGTGGGC 91  
|||||  
QY 75 GCGGGCCGATGGCGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAGAAAGTGAAGCT 134  
|||||  
Db 92 GCGGGCCGATGGCGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAGAAAGTGAAGCT 151  
|||||  
QY 135 CAGCACATCTGCGCCAGTCCACACAGACAGAACCTATGTCACTTTCAAGCTCTGGGT 194  
|||||  
Db 152 CAGCACATCTGCGCCAGTGCACCAACAGACAGAACTATGTCACTTTCAAGCTCTGGGT 211  
|||||  
QY 195 TGCCGAGCACACGACTCCAAACAGAGAGTGTGTCAGAGCTGCTCGTCACTCTTTGGC 254  
|||||  
Db 212 TGCCGAGCACACGACTCCAAACAGAGAGTGTGTCAGAGCTGCTCGTCACTCTTTGGC 271  
|||||  
QY 255 ACCAAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGTGGCTCTGTGTGTCACCACTGAA 314  
|||||  
Db 272 ACCAAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGTGGCTCTGTGTGTCACCACTGAA 331  
|||||  
QY 315 CACATCTTGGTTCGCTGGGCTGCTGGGGTGTCTGTAGCACCTTGAGGAGCTGCTGCC 374  
|||||  
Db 332 CACATCTTGGTTCGCTGGGCTGCTGGGGTGTCTGTAGCACCTTGAGGAGCTGCTGCC 391  
|||||  
QY 375 CCAACACACGCGTGTGCGGTCTTGGCCAACTGCGCTGTGTGTCAGGAGGGGCC 434  
|||||  
Db 392 CCAACACACGCGTGTGCGGTCTTGGCCAACTGCGCTGTGTGTCAGGAGGGGCC 451  
|||||  
QY 435 ATAGTGAATGGCGGGGCCGCCACGCTGGGAGCAGCAGACCAAGCTTCCTGCAGCATCTG 494  
|||||  
Db 452 ATAGTGAATGGCGGGGCCGCCACGCTGGGAGCAGCAGACCAAGCTTCCTGCAGCATCTG 511  
|||||  
QY 495 CTGGAGGCTGTGGGGGTGTGAGAGGTGCCTGAAGCCTAGCTCGACATCCACACTGGC 554  
|||||  
Db 512 CTGGAGGCTGTGGGGGTGTGAGAGGTGCCTGAAGCCTAGCTCGACATCCACACTGGC 571  
|||||  
QY 555 CTCAGTGGCAGTGGCGCTTGTGTGATTCCTCGAGGCGCTTGCTCAAGAGGCC 614  
|||||  
Db 572 CTCAGTGGCAGTGGCGCTTGTGTGATTCCTCGAGGCGCTTGCTCAAGAGGCC 631  
|||||  
QY 615 GTCAAGATGGGATGCCAGCAGCTTGCCGCCACCGCATCGTGCACACCTGCTGGGG 674  
|||||  
Db 632 GTCAAGATGGGATGCCAGCAGCTTGCCGCCACCGCATCGTGCACACCTGCTGGGG 691  
|||||  
QY 675 ACGGCCAAGATGCTGTGCAGAGGGCCAAACCCAGCCAGCTGGCTCAGACCTGTGTC 734  
|||||  
Db 692 ACGGCCAAGATGCTGTGCAGAGGGCCAAACCCAGCCAGCTGGCTCAGACCTGTGTC 751  
|||||  
QY 735 ACCCGGGTGGCACCACCATCTATGAGCTCCACGCCCTGGAGCAGGGGGCTGGAGCA 794  
|||||  
Db 752 ACCCGGGTGGCACCACCATCTATGAGCTCCACGCCCTGGAGCAGGGGGCTGGAGCA 811  
|||||  
QY 795 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCCGGGCCAAGAGAGCTCAGCAGAAAGTAG 854  
|||||  
Db 812 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCCGGGCCAAGAGAGCTCAGCAGAAAGTAG 871  
|||||  
QY 855 GCTGGGCTGTGGCCATTCCTTCTGCTGTGCGCCCTGCGCTTCCTGTGTCCTTCCC 914  
|||||  
Db 872 GCTGGGCTGTGGCCATTCCTTCTGCTGTGCGCCCTGCGCTTCCTGTGTCCTTCCC 931  
|||||  
QY 915 CTGAGGACTGCGGCTCCCTCCCTCTGATGAGGCTCTCCCTACTGCTCTTCTCCCTTTG 974  
|||||  
Db 932 CTGAGGACTGCGGCTCCCTCCCTCTGATGAGGCTCTCCCTACTGCTCTTCTTCCCTTTG 991  
|||||  
QY 975 CACAGGAAATGCAGGGGGCAGGACTTGGAGGTTCCAGAGCGCGGGGAGCCCGACCA 1034  
|||||  
Db 992 CACAGGAAATGCAGGGGGCAGGACTTGGAGGTTCCAGAGCGCGGGGAGCCCGACCA 1051  
|||||  
QY 1035 GTGGGAGACACTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTTT 1094  
|||||  
Db 1052 GTGGGAGACACTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTTT 1111  
|||||







useful in diagnosis and gene therapy -

Claim 1; Page 1961-1966; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAW79323-AAW80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

Sequence 5055 BP; 989 A; 1736 C; 1512 G; 818 T; 0 other;

Query Match	54.0%;	Score 723;	DB 22;	Length 5055;
-------------	--------	------------	--------	--------------

Best Local Similarity 95.9%; Pred. No. 9.2e-152;

<b>Matches</b>	<b>750;</b>	<b>Conservative</b>	<b>0;</b>	<b>Mismatches</b>	<b>30;</b>	<b>Indels</b>	<b>2;</b>	<b>Gaps</b>	<b>1;</b>
----------------	-------------	---------------------	-----------	-------------------	------------	---------------	-----------	-------------	-----------

QY 45 CCGTCTCCGGCGCGCTGGGCTTCGTGGCGCGGGCCGCATGGCGGGGCCATCGCGCAG 104

100

QV 105 GGCCTCATCAGAGCAGGAAAGTGGAAAGCTCAGCACATACTGGCCAGTGCACCAACAGAC 164

[illegible]

05 AGGAACCTATGTCACCTTTCAAGCTCTGGGTTGGGACCAAGCACTCCACAGGACGTC 224

Q7 103 AUGAACCAATAGTCACATTTCAAGCTCIGGGTGTGCGGACCAACGCACCTCCATACCAAGGAGGTG 224

[illegible]

QY  
Z23 CTGCAGAGCTGCGCTGCTGCTCATCTTTGCCACCAAGCCCTCATGTGCTGCCAGCTGTCTCTG 284

285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246

QY 283 GCAGAGTGGCTCCGTGGTGCACCACTGAACACATCTTGGTGTCCCGTGGCTGGCTGGGGTG 344

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

345 TCCTGAGCACCCCTGGAGGAGCTGCTGCCCCCAACACACGGGTGCTGGGGTCTTGCCC 404  
QY

405 AACCTGCCCTGTGTGGTCCAGGAAGGGCCATAGTGATGGCGGGGGCCACGTGGGG 464

.....

Qy 465 AGCAGCGAGACCAAGCTCCTGCAGCATCTGCTGGAGGCCCTGTGGCGGTGTGAGGAGGTG 524

.....

QY 525 CCTGAAGCCTACGTGACATCCACACTGGCCTCAGTGGCAGTGGCGTGGCCTTCGTGTGT 584

TOTAL

QY 585 GCATTCTCCGAGGCCCTGGCTGAAGGAGCCGTCAGATGGGCATGCCAGCCTGGCC 644

[illegible]

QY 645 CACCGCATCGCTGCCCAGACCCCTGCTGGGACGGCCAAGATGCTGCTGCACGAGGGCAA 704

[illegible]

QY 705 CACCCAGCCAGCTGGCTCAGACGTGTGCACCCCGGTGGCACCACCATCTATGGACTC 764







PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	11-DEC-2000;	2000US-0251990.
PR	05-JAN-2001;	2000US-0254097.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA		
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-465566/50.	
DR	P-PSDB; AAU23472.	
XX		
PT	Novel polypeptides and polynucleotides useful for diagnosing,	
PT	preventing, treating neural, immune system, muscular, reproductive,	
PT	pulmonary, cardiovascular, renal, proliferative disorders and cancerous	
PT	diseases	
XX		
XX	Claim 4; SEQ ID NO 568; 1180pp; English.	
XX		
CC	The present invention relates to the isolation of novel human enzyme	
CC	polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences	
CC	encoding them. The enzyme polypeptides of the invention may comprise the	
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	
CC	isomerases or ligases. The sequences of the invention are useful in the	
CC	diagnosis, treatment, prevention and/or prognosis of a wide range of	
CC	disorders including hyperproliferative disorders (e.g. cancer),	
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders	
CC	(e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),	
CC	metabolic disorders (e.g. phenylketonuria), inflammatory disorders	
CC	(e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),	
CC	blood-related disorders (e.g. haemophilia), reproductive disorders	
CC	(e.g. infertility) and infectious disorders (e.g. influenza). The	
CC	polynucleotides of the invention can also be used in gene therapy.	
CC	AAS40785-AAS41084 represent cDNA sequences encoding for the novel human	
CC	enzyme polypeptides of the invention.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at <a href="http://ftp.wipo.int/pub/published_pct_sequences">ftp.wipo.int/pub/published_pct_sequences</a> .	
XX		
SQ	Sequence 270 BP; 54 A; 82 C; 88 G; 45 T; 1 other;	
	Query Match 18.1%; Score 241.8; DB 22; Length 270;	
	Best Local Similarity 95.9%; Pred. No. 1.3e-44;	
	Matches 258; Conservative 1; Mismatches 8; Indels 2; Gaps 1;	
QY	10 AGTCTGTCGGAGCAACAAGATGGCAGCTGCAGCGCTCCGCCGGCGGTGGCTTCG 69	
Db	2 AGGCGCTCGAGGACACAGAATGGCAGCTTCGCGGAGCCGCTCCGCCGGCGGTGGCTTCG 61	
QY	70 TGGGCGCGGGCCGCATGGCGGGGGCCCATCCGCGAGGGCCCTCATCAGACGAGGAAGTGG 129	
Db	62 TGGGCGCGGGCCGCATGGSGGGGGCCCATCCGCGAGGGCCCTCATCAGACGAGGAAGTGG 121	
QY	130 AAGCTCAGCACATCTGGCCAGTGCCACCACACAGACGAGAACCTATCTCACTTTCAAGTTC 189	
Db	122 AAGCTCAGCACATCTGGCCAGTGCCACCACACAGACGAGAACCTATCTCACTTTCAAGTTC 181	
QY	190 TGGGTTTGGCGGACACGCACTGCCAACCAGGAGGTGCTGCAGAGCT--GCCTGCTGTCAT 247	

Db	182	TGGGTTGGCGGACCAAGCAGCTCAACACGAGAGTGTGCAGAGCTTGCCTGCTTCAT	24
QY	248	CTTTGGCCACCAAGCCTCATGTGCTGCCAG	276
Db	242	CTTTGGCCACCAAGCCTCATGTGCTGCCAG	270
RESULT 13			
ID	ABA00760	standard; DNA; 1848 BP.	
XX	ABA00760;		
XX	18-MAR-2003	(first entry)	
DE	XX	P5CR related nucleic acid #3.	
KW	XX	Pyroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;	
KW	XX	pyroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary; ss.	
OS	XX	Homo sapiens.	
FN	XX	WO200299043-A2.	
PD	XX	12-DEC-2002.	
XX	XX	03-JUN-2002; 2002WO-US17319.	
XX	XX	05-JUN-2001; 2001US-296080P.	
PR	FR	10-OCT-2001; 2001US-328509P.	
XX	XX	(EXEL-) EXELIS INC.	
XX	XX	Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;	
PI	Engst S;		
XX	XX	WPI; 2003-140606/13.	
DR	XX	Identifying p53 pathway modulating agents with P5CR genes, useful for	
PT	the diagnosis and treatment of disorders associated with defects in the		
PT	p53 pathway, such as cancer of the breast, colon, kidneys, lung and		
PT	ovary		
XX	Disclosure; Page 40-41; 53pp; English.		
XX	XX	The sequences given in ABA00758-66 represent pyrroline 5 carboxylate	
CC	reductase (P5CR) related nucleic acid sequences. P5CR catalyses the		
CC	NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to		
CC	proline. These sequences may be used in the method of the invention		
CC	for identifying a candidate p53 pathway modulating agent. The method		
CC	comprises providing an assay system comprising a purified P5CR		
CC	polypeptide or nucleic acid, or a functionally active fragment or		
CC	derivative, contacting the assay system with a test agent, where the		
CC	system provides a reference activity, and detecting a test agent-biased		
CC	activity of the assay system. The method of the present invention is		
CC	useful for the diagnosis and treatment of disorders associated with		
CC	defects in the p53 pathway, such as cancer of the breast, colon,		
CC	kidneys, lung and ovary.		
SQ	Sequence 1848 BP; 364 A; 582 C; 541 G; 361 T; 0 other;		
	Query Match 16.9%; Score 226.4; DB 25; Length 1848;		
	Best Local Similarity 58.1%; Pred. No. 5.3e-41;		
	Matches 467; Conservative 0; Mismatches 316; Ind		





Db	551	CGCATTCACGCCCTGGATGCCCTGCTGATGGGGGTGTCAGATGGGACTTCCAAAGCGC	610
QY	638	CCTGGCCCAACGGCATCGCTGCCAGACCCCTGCTGGGAGCGGCCAAGATGCTCTCTGCACGA	697
Db	611	CCTGCGAGTGGCTCGGGGCCACAGGCCCTTCCTGGGGGCTGCCAAGATGCTCTGCACTC	670
QY	698	GGGCCAACACCCAGCCAGCTGGGCTCAGACAGTGTGCACCCGGGTGGGCACCACCATCTA	757
Db	671	AGACAGACACCCAGGCCAGCTCAAGGACACGTCAGCTCTCTGTGGGGCCACCACCTCA	730
QY	758	TGGACTCCAGCCCTGGAGCAGGGCGGGCTGCAGCAGCCACCATCAGCGCGTGGAGGC	817
Db	731	TGCCCTTGATGTGTGGAGAGTGGGGCTTCCGCTCCCTGCTCATCAACGCTGTGGAGGC	790
QY	818	TGCCACCTGCCGGGCGCAAGAGCT	841
b	791	CTCTTGCAATCGGCACACGGGAGCT	814

RESULT 15

ABN95903  
ID ABN95903 standard; DNA; 1792 BP.

AC ABN95903;

DT 13-AUG-2002 (first entry)

DE Gene #2401 used to diagnose liver cancer.

Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic; metastatic liver tumour; cytostatic; expression profile; disease state; disease progression; drug toxicity; drug efficacy; drug metabolism.

AA  
OS Homo sapiens.

AA WO200229103-A2.  
PN

11-APR-2002.

XX  
PF  
02-OCT-2001: 2001WO-US30589.XX  
PR 02-OCT-2000: 2000US-237054P.

PA (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG; PT

XX  
DR WPT: 2002-426119/45.

xx Diagnosing and detecting the progression of liver cancer,  
 pt hepatocellular carcinoma or metastatic liver tumor in a patient.  
 pt Involves detecting the level of expression of two or more genes in a  
 pt liver tissue sample -

PS Claim 1: SEO ID NO 2401; 298pp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytotactic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

CC drug toxicity, drug efficacy and drug metabolism.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

[illegible]

Search completed: August 8, 2003, 16:48:57  
Job time : 403 secs